Education
L
11
.J8525
SCHOOL OF EDUCATION
LIBRARY

The Journal of Experimental Education

A periodical report of scientific investigations relating to child development, curriculum, learning, teaching, supervision, measurements, statistics, and experimental techniques.

Volume XII | SEPTEMBER, 1943 | Number 1

LEARNING, TEACHING, SUPERVISION

CONTENTS

PAGE

The Effects of Utilizing Selected Community Resources in Ninth-Grade and Tenth-Grade Homemaking Classes: *Edna D. Meshke* ---------- 1

An Experimental Evaluation of the School Excursion: *Ella Callista Clark*---------- 10

Comparative Effectiveness of Pictorial Teaching Aids: *David J. Goodman*---------- 20

Criteria For Evaluating Reading Textbooks: *David Kopel and John F. O'Connor* 26

Procedures For Evaluating Textbooks in Reading: *David Kopel* ---------- 34

Reading Reactions For Varied Types of Subject Matter: *Earl William Seibert*---- 37

The Plot–Completion Test: *Sarah I. Roody* ---------- 45

Influence of Student Dreads Upon Attitudes Toward School Subjects: *Winona M. Perry* ---------- 48

$5.00 A YEAR | PUBLISHED QUARTERLY | $1.50 A COPY

Edited and Published by A. S. Barr, Professor of Education, University of Wisconsin, Madison, Wisconsin.

Entered as second-class matter October 17, 1938 at the post office at Madison, Wisconsin under the act of March 3, 1879.

EDITORIAL BOARD

A. S. Barr, Chairman, Professor of Education, University of Wisconsin, Madison, Wis.

Carter V. Good, Professor of Education, University of Cincinnati, Cincinnati, Ohio. Editorially responsible for materials on learning, teaching and supervision, published each September.

J. Wayne Wrightstone, Assistant Director, Bureau of Reference, Research and Statistics, Board of Education of the City of New York, 110 Livingston Street, Brooklyn, New York. Editorially responsible for materials on curriculum construction, published each June.

Edward E. Cureton, Senior Personnel Technician, Personnel Research and Occupational Analysis Section, The Adjutant General's Office, 270 Madison Avenue, New York 16, N. Y. Editorially responsible for materials on measurements, statistics, and methods of experimental research, published each March.

Arthur T. Jersild, Professor of Education, Teachers College, Columbia University, New York City. Editorially responsible for materials on child welfare, guidance, and development, published each December.

CONTRIBUTING EDITORS

Gilbert L. Betts, Supervisor of Graduate Research in Education, Colorado State College, Fort Collins, Colorado.

William A. Brownell, Professor of Educational Psychology, Duke University, Durham, North Carolina.

Leo J. Brueckner, Professor of Education, University of Minnesota, Minneapolis, Minnesota.

Oscar K. Buros, Associate Professor of Education, Rutgers University, New Brunswick, New Jersey.

Otis W. Caldwell, General Secretary, The American Association for the Advancement of Science, Boyce Thompson Institute for Plant Research, Yonkers, New York.

Leslie L. Chisholm, Associate Professor of Education, State College of Washington, Pullman, Washington.

Herbert S. Conrad, College Entrance Examination Board, Princeton, N. J.

Stephen M. Corey, Professor of Educational Psychology, University of Chicago, Chicago, Illinois.

Robert A. Davis, Professor of Education, Director of Bureau of Educational Research, University of Colorado, Boulder, Colorado.

Harl R. Douglass, Director of College of Education, University of Colorado, Boulder, Colorado.

Jack W. Dunlap, Associate Professor of Educational Psychology, University of Rochester, Rochester, New York.

Harold A. Edgerton, Assistant Professor of Psychology, Ohio State University, Columbus, Ohio.

Lt. Comdr., A. C. Eurich, USNR, Officer in charge, Standards and Curriculum, Bureau of Naval Personnel, USN, 3130 Wisconsin Avenue, N. W., Washington 15, D. C.

Lt. Col. John C. Flanagen, Chief, Psychological Branch, Office of Air Surgeon, Headquarters, Army Air Forces, Washington 25, D. C.

Kai Jensen, Professor of Education, University of Wisconsin, Madison, Wisconsin.

Harold E. Jones, Professor of Psychology and Director, Institute of Child Welfare, University of California, Berkeley 4, California.

Noel Keys, Professor of Education and Lecturer in Human Relations, University of California, Berkeley, California.

Edward A. Lincoln, Consulting Psychologist, Halifax, Massachusetts.

T. E. Newland, Lt., USNR, 2702 Wisconsin Avenue, N. W., Washington 7, D. C.

C. W. Odell, Associate Professor of Education, University of Illinois, Urbana, Illinois.

Willard C. Olson, Professor of Education, Director of Research in Child Development, University of Michigan, Ann Arbor, Michigan.

W. E. Peik, Dean and Professor of Education, University of Minnesota, Minneapolis, Minnesota.

S. L. Pressey, Professor of Educational Psychology, Ohio State University, Columbus, Ohio.

Clarence E. Ragsdale, Professor of Education, University of Wisconsin, Madison, Wisconsin.

H. H. Remmers, Director, Division of Educational Reference, Professor of Education and Psychology, Purdue University, Lafayette, Indiana.

Henry D. Rinsland, Professor of Education and Director of Educational Research, The University of Oklahoma, Norman, Oklahoma.

Robert T. Rock, Jr., Professor of Psychology, Head of Dept. of Psychology, Graduate School, Fordham University, New York City.

G. M. Ruch, Chief of Research and Statistical Service, U. S. Office of Education, Washington 25, D. C.

Phillip J. Rulon, Associate Professor of Education, Harvard University, Cambridge 38, Massachusetts.

Douglas E. Scates, Associate Professor of Education, Duke University, Durham, North Carolina.

David Segel, Educational Consultant, Specialist in Tests and Measurements, Federal Security Agency, U. S. Office of Education, Washington, D. C.

Paul W. Terry, Professor of Educational Psychology, University of Alabama, University, Alabama.

Helen Thompson, Clinic of Child Development, Research Associate, Yale University, New Haven, Connecticut.

Robert L. Thorndike, Associate Professor of Education, Teachers College, Columbia University, New York City.

Herbert A. Toops, Professor of Psychology, Ohio State University, Columbus, Ohio.

T. L. Torgerson, Professor of Education, University of Wisconsin, Madison, Wisconsin.

Helen M. Walker, Professor of Education, Teachers College, Columbia University, New York City.

Beth L. Wellman, Professor of Psychology, Child Welfare Research Station, State University of Iowa, Iowa City, Iowa.

Guy M. Wilson, Professor of Education, Boston University, Boston, Massachusetts.

Paul A. Witty, Professor of Education, Director of Psycho-Educational Clinic, School of Education, Northwestern University, Evanston, Illinois.

Ernest R. Wood, Professor of Education, New York University, New York City.

DEMOCRAT PRINTING COMPANY
MADISON, WISCONSIN

J

Vol

TH

T
to
of
and
use
stru
who
mu
acti
eith
crea
resu
wer

T
exte
yea
and
with
tent
220
trod
four
C
tive
serv
offe
tent
reas
E
Sele
a p
nari
clas
four
and
class
E
"Sel

[1] Al
Unive
Clara

Journal of Experimental Education

Volume XII | September, 1943 | Number 1

THE EFFECTS OF UTILIZING SELECTED COMMUNITY RESOURCES IN NINTH-GRADE AND TENTH-GRADE HOMEMAKING CLASSES[1]

Edna D. Meshke
Butler University

The present investigation was undertaken to discover: (1) whether, within the structure of the typical secondary-school curriculum and schedules, community resources could be used in connection with home economics instruction; (2) the effectiveness of instruction when it included direct contacts in the community and when it was limited to classroom activities; and (3) whether contacts with either type of experimental procedure increased pupil achievement over that which resulted from instruction when such contacts were lacking.

Scope of the Problem

The study began in February, 1941, and extended over a period of approximately one year. It was carried on in ten communities and included fourteen ninth-grade classes with an enrollment of 268 girls, and thirteen tenth-grade classes with an enrollment of 220 girls. Experimental procedures were introduced into seven schools, and the other four were used as control schools.

Communities were selected that had relatively similar types of store and electrical services and in which home economics was offered at both the ninth-grade and the tenth-grade levels and in departments with reasonably adequate space and equipment.

Eight ninth-grade classes studied "Food Selection and Purchase" for twelve weeks as a part of the regular food study unit ordinarily taught during the fall semester. These classes were taught by five different teachers; four of them were taught with store contacts and four with instruction limited to the classroom.

Eight tenth-grade classes studied the "Selection and Care of Electrical Equipment" for four consecutive weeks as a part of the house furnishing and equipment unit ordinarily included in that grade. These classes were taught by five different teachers; four of them were taught with store contacts and four with instruction limited to the classroom.

In addition, ten control classes were used: five in the ninth grade and five in the tenth grade.

[1] Abstract of a doctorate dissertation completed at the University of Minnesota in 1942, under the direction of Clara M. Brown.

Procedure

The tentative selection of the communities was made in March, 1941, after an analysis of questionnaire returns and following a visit to the high schools.

Instructional materials and evaluation devices were prepared during the summer and early fall; the schools were revisited and a half-day conference was held with the participating teachers during September; and the experimental procedures were inaugurated in most schools early in October. The last data were collected in March, 1942.

The final decision as to which schools would be used for experimental and which for control classes was determined by the number of ninth- and tenth-grade classes in the school and its accessibility to Minneapolis and to other communities which would be used. The latter criterion was important because, during the study, the author planned to make several visits to the experimental schools and only one visit to each control school.

Instructional Procedures Used With Different Groups

The ninth-grade and the tenth-grade classes were divided into three groups: those classes in which pupils had direct contact with the stores and referred to hereafter as the "store" group; those classes in which

pupils had contact only with classroom materials and referred to as the "classroom" group; and those in which no experimental procedures were used and referred to as the "control" group.

"Store classes" took one exploratory excursion with the teacher at the beginning of the unit to become acquainted with the location and general characteristics of the local stores; and this was followed by indivdual and group trips taken by pupils during class time. In the ninth grade, each pupil had two opportunities to purchase supplies for a particular foods lesson or for use in the department for other classes. In the tenth grade, six field trips were taken during the four-week unit to investigate the six pieces of electrical equipment studied.

"Classroom classes" used a variety of descriptive and pictorial materials to supplement class discussions. In the ninth grade, the teachers supplied the classes with weekly sales sheets from the local grocery stores and with food advertisements from newspapers; these furnished the basic information on prices, qualities, sizes, and the foods that were in season. In the tenth grade, the teachers collected and made available to the pupils illustrated booklets, pamphets, and other supplementary descriptive materials dealing with electrical equipment.

"Control classes" were taught in any way that the teacher desired. No explanation was made of the plans developed for use in either of the experimental groups, and the teachers were permitted to teach the units for the length of time they considered desirable for their pupils. The only requirements made of the teachers were that they teach units in food selection and purchase, and in the selection and care of electrical equipment, and that they administer the necessary tests which the investigator supplied before instruction and at the close of the units.

Development of Instructional Materials and Techniques of Evaluation

There seemed to be little need for developing any special instructional materials for the ninth-grade foods unit, because there were texts, references, and other materials easily available; but a very different situation existed in the case of the tenth-grade unit on electrical equipment. Therefore, the materials developed for each grade are discussed separately.

Ninth Grade.—Tentative objectives for instruction in food selection and purchase were formulated on the basis of previous knowledge of ninth-grade home economics work. It was known that pupils were frequently sent to the stores to buy groceries for home use, and it was assumed that they had certain attitudes and encountered certain problems in connection with their job as buyers, such as how much to buy, where to buy, how much to pay, what to do if the price was too high, how to judge quality, and how to do the buying in the quickest way and with the greatest satisfaction.

The pencil-and-paper test covered such problems as those cited. The test items were constructed after a study of high school texts, reference books, food advertisements, and weekly sales sheets from the stores; and visits to grocery stores to examine labels and to discover features that might attract the average shopper.

Check lists and rating devices were developed for the pupils to record their behavior when buying, the decisions they would make under certain described conditions, and their judgment concerning groceries bought by someone else. A list of general suggestions for the food buyer, and three short self-administering tests, were formulated for use as study guides for both store and classroom classes.

Tenth Grade.—Pupil goals for the study of electrical equipment were believed to be less tangible and personal than were those for the unit on foods. However, it was assumed that those who had not studied a body of content dealing specifically with electrical equipment were likely to be interested in learning how much it costs to use electrical equipment, how a "good buy" can be recognized, what a guarantee really means, how long a piece of equipment will last, and other information of a similar practical nature.

A survey of available library materials dealing with some aspect of electrical equipment and of the illustrative and descriptive booklets from manufacturers led to the conclusion that a great deal of material was available, ranging from fragmentary technical discussions to broad generalities, and an inspection of many pieces of equipment in the stores showed that much information could be

School of Educ. Lib
Educ
Direct
2-3-47
Heritage
v. 12, no. 1

secured from dealers and from a careful examination of merchandise. However, it seemed reasonable to conclude that high school teachers would not have the time, the patience, or the desire to canvass all of these sources to secure instructional materials for a short unit. Also, it seemed desirable that the pupils in both the store and the classroom groups in the tenth grade should have access to at least some of the same instructional materials. Consequently, prior to the inauguration of the experiment, a 44-page booklet was prepared covering the basic facts pertaining to the six pieces of electrical equipment that were most frequently available in homes,[2] according to a recent survey. Characteristic features of each piece of equipment were described for low, medium, and high price levels; and definite suggestions were presented to serve as guides for the selection and care of the equipment.

To establish a basis for uniformity of emphasis in the experimental classes, it seemed desirable to provide guides for discussing the factual material and its applications. For this purpose, the home situations of three families were described, and an explanation made of their needs for new equipment and of the problems involved before a purchase could be made.

The pencil-and-paper test consisted of judgment questions based upon the factual content of the booklet and upon the statements found in advertisements of electrical equipment.

Cooperative Development of Experimental Procedures

Early in September, 1941, seven schools, selected for the experimental situations, were re-visited and a joint planning conference was proposed to the home economics teachers.

The teachers came to the university campus one Saturday in September, at their own expense, for a half-day meeting to discuss goals, content, source materials, pupil activities, and possibilities for checking outcomes. The investigator announced no predetermined plans for the experimental procedures at this conference. Instead, she worked out with the teachers the pupil goals for the units, in terms of the problems that their pupils were concerned with when selecting food, and in the light of the questions which they believed their pupils would ask about electrical equipment. In addition, the cooperating teachers discussed the content of each unit and listed the teaching procedures that they believed would be effective in attaining the goals set up and would be appropriate to the subject matter of the particular unit; and a careful analysis was made of the differences between the two experimental procedures in terms of materials of instruction, pupil activities, and means of evaluation.

Descriptions were presented of the instructional materials in the process of construction, and the teachers were told that such materials would be made available to them and to their pupils when each unit was begun in their schools.

It was announced that the experimental procedure followed in each class would be determined by chance and that on the day of the first pre-test the specific assignment would be made. Teachers indicated tentatively when they would be ready to begin the experimental units in each grade. The exact dates for the administration of pre-tests were determined later by correspondence.

Both "store" and "classroom" procedures were used in each school where ninth- and tenth-grade classes were taught, and the assignment of procedures to specific classes was made without definite knowledge about the pupils or of the teacher's effectiveness, but chiefly from the standpoint of securing an approximately equal number of classes in which each experimental procedure would be followed.

Investigator's Contacts With Schools

The investigator administered the pretest for each unit in each experimental school when the teacher indicated that she was ready to introduce the experimental procedures. Assistance was given to enable the teachers to follow the specified procedures in connection with regular school routines and to interpret the use of the instructional materials and check lists. One or more visits were made to each experimental school during the progress of each unit, and again at the close of the experimental period, in order to administer the re-test and to collect all pupil records.

Visits were made to the control schools for the purpose of administering the pretests in

[2] Edna Meshke, *Selection and Care of Electrical Equipment*. Minneapolis: Burgess Publishing Company, 1941.

each grade, but the teachers in these schools administered the tests themselves when the work had been completed.

Analysis of Data

The following data were collected: the pretest score, the retest score (the same form of the pencil-and-paper test was used for both the pretest and the retest in each grade), the IQ based on the Otis test,[2] and the socio-economic status based upon the father's occupation. It was found that the father's occupation had relatively little influence upon the daughter's retest score, since the correlation was negligible; hence, no consideration was given to occupational status in the statistical analysis of the retest scores.

The coefficient of reliability of the pencil-and-paper tests, administered as retests, was .90 for the ninth-grade test and .87 for the tenth-grade test. The split-score method was used to compute the reliability of the tests, and the resulting coefficients were stepped up by application of the Spearman–Brown formula.

The check-list ratings and the pupil comments that were collected as supplementary evidence were not subjected to rigorous statistical analysis, but the test data were carefully analyzed. The statistical treatment of the test data involved primarily two techniques, namely, the analysis of variance and the analysis of covariance.

The analysis of variance was used to discover whether the several classes were sufficiently homogeneous with respect to their means to be considered as representative samples from the same population.

The technique used to test for homogeneity of means makes it possible to isolate and compare the variations derived from two sources: *within* each class and *among* the classes. The formulas used in computing these estimates of variation are based upon the principle of least squares which, when developed algebraically, shows that the mean is the value from which each individual score deviates the least when the variation is measured by squaring the deviations.

The estimates of variation both within classes and among classes, expressed as mean squares, are quotients in which the sum of the squares represents the numerator and the degrees of freedom or number of independent comparisons that it is possible to make represents the denominator of the value.

The F-ratios, obtained by comparing the mean square *among* a number of classes with the mean square *within* these classes, were checked with tabeled values of F at the .05 level and at the .01 level for the appropriate degrees of freedom, to discover whether the differences were great enough to be significant or whether they were likely to be merely chance occurrences.

The .01 value of F was selected as the acceptable level of significance in the present study, and all observed F-values greater than the one percent point were considered statistically significant or, in other words, indicated that the groups being compared were not homogeneous.

In order that the results obtained in the present study might have predictive value, it was necessary to discover the "consistency in outcomes" among a number of classes in which a specified teaching procedure had been used, and to base conclusions regarding the effectiveness of the procedure upon the results observed from as many replications as possible. Therefore, the retest scores were used as the criterion for determining whether the classes were sufficiently homogeneous to be considered as samples representative of the same population. If homogeneity of retest scores was demonstrated, heterogeneity of means of the IQ's and initial scores were not considered in grouping classes, because adjustments for differences in initial measures were to be made by covariance in all analyses in which groups of classes that were homogeneous on retest scores were treated as a single sample.

Analysis of covariance involves the analysis of the sum of cross-products of deviations into component parts in the same way as the sum of squares of deviations is analyzed into component parts when only one variable is considered. The actual adjustment of the retest scores in the present study was made in terms of the multiple regression equation between the retest scores and the IQ's and the initial test scores.

It was necessary to test the basic assumptions underlying the use of the analysis of variance (normality of distribution and homogeneity of variance) and the analysis of covariance (homogeneity of regression and homogeneity of residual variance) before one

[2] Quick Scoring; Form Beta A for the ninth grade and Form Gamma C for the tenth grade.

TABLE I

MEANS OF IQ, PRETEST SCORES, AND RETEST SCORES FOR THE NINTH-GRADE CLASSES

Schools	IQ	Mean Scores On Pretest	Mean Scores On Retest	Mean Gains
Store Group				
A	107.71	59.96	74.64	14.68
B	99.50	46.31	55.81	9.50
C	99.67	51.33	61.33	10.00
Classroom Group				
A	98.67	49.48	59.04	9.56
B	102.87	48.60	52.87	4.27
C	100.22	51.17	54.61	3.44
D	108.50	50.22	55.78	5.56
E	106.15	49.30	68.65	19.35
Control Group				
G_1	100.11	44.16	46.47	2.31
G_2	103.50	46.05	50.64	4.59
H	109.70	47.91	50.48	2.57
I	109.16	52.32	55.84	3.52
J	103.36	53.18	57.59	4.41

was justified in grouping classes shown to be homogeneous with respect to means of the retest scores or the adjusted retest scores.

Since the F-test permits fairly wide departure from normality and since inspection of the data gave no indication of such departures, it was deemed unnecessary to test the assumption of normal distribution. However, the other basic assumptions were tested and satisfactorily demonstrated in each situation where the analysis of variance and the analysis of covariance were used in the present study.

In each group of ninth-grade classes (store, classroom, and control) one class had to be eliminated in order to have groups that were homogeneous with respect to the retest scores. School A had to be eliminated from the store group, School E from the classroom, and either School J or the poorer class from School G had to be eliminated from the control group. Reference to Table I will show how much each of the ninth-grade classes that had to be eliminated differed from the others in the group.

In the tenth grade the four classes comprising the store group were homogeneous with respect to retest means, which was likewise true of the four classes in the classroom group (See Table II). However, in the con-

TABLE II

MEANS OF IQ, PRETEST SCORES, AND RETEST SCORES FOR THE TENTH-GRADE CLASSES

Schools	IQ	Mean Scores On Pretest	Mean Scores On Retest	Mean Gains
Store Group				
A	103.61	49.48	69.70	20.22
B	99.39	50.48	76.61	26.13
D	104.63	53.60	81.88	28.28
E	96.65	49.04	71.35	22.31
Classroom Group				
A	103.96	47.08	69.16	22.08
B	100.82	49.55	71.95	22.40
C	91.10	43.90	58.30	14.40
F	102.47	51.82	72.06	20.24
Control Group				
G_1	106.00	53.77	64.92	11.15
G_2	100.75	51.25	61.58	10.33
H	105.83	51.17	53.17	2.00
I	106.13	51.75	52.75	1.00
J	106.38	45.19	72.81	27.62

trol group, the tenth-grade class from School J had to be eliminated in order to establish homogeneity, and it will be noted that the achievement of the pupils in this school was comparable with that of pupils who had store contacts.[4]

Findings

The difference in the achievement of the pupils in experimental classes with and without direct store contacts during class time was consistently in favor of those who had the contacts with the stores, both when buying groceries and when investigating electrical equipment. The statistical analyses of the results of the pencil-and-paper test are reported separately for the ninth and the tenth grades.

Ninth Grade.—The difference between store and classroom groups was in favor of the classes in which pupils had direct store contacts, whether comparisons were made between classes taught by the same teacher or by different teachers.

Table III shows that there were significant differences between the mean achievement of the store and the classroom groups, since the F-ratios for the adjusted retest scores were 8.30 and 9.98 in contrast to the tabeled F-ratios, at the .01 level, of 7.04 and 6.96, respectively. It will be noted that in the first comparison classes were included from only two schools in which both store and classroom procedures were employed, and in the second comparison an additional school was added in which only the classroom procedures were used in the ninth-grade class.

Scrutiny of Table I will show that the store group made a higher mean achievement than did the classroom group; and, therefore, the significant F-ratios reported in Table III are in favor of the pupils who had the direct contacts with the stores.

Further study of Table I will show that the classroom group gained slightly more than did the control group, and the store group gained considerably more than did the control group on the pencil-and-paper test. However, Table III shows that only in the comparison of the store and the control groups were the differences in achievement statistically significant.

In the comparisons of the classroom and the control groups, the differences were not large enough to reach the tabled F-values, even at the .05 level. On the other hand, the differences in favor of the store group, when compared with the control groups, far exceeded the tabled F-values at the .01 level, being 30.12 and 20.12 when the tabled values were 6.87 and 6.84, respectively.

The relative gains made by classes with and without store contacts, when both procedures were used by the same teacher, are presented in Table IV in terms of standard deviations.

[4] It was discovered that the booklet, *Selection and Care of Electrical Equipment*, intended for use in experimental classes only had been used as a basic text in School J.

Table III

F-Ratios of Adjusted Retest Scores of Ninth-Grade Experimental and Control Groups Drawn from Seven Schools

Groups and Schools	Theoretical F-Ratios* (.05)	(.01)	Observed F-Ratios
Store and Classroom			
Store (B, C) / Classroom (B, C)	3.99	7.04	8.30
Store (B, C) / Classroom (B, C, D)	3.96	6.96	9.98
Classroom and Control			
Classroom (B, C, D) / Control ($G_{1, 2}$, H, I)	3.92	6.48	3.40
Classroom (B, C, D) / Control (G_1, H, I, J)	3.92	6.84	Less than 1
Store and Control			
Store (B, C) / Control ($G_{1, 2}$, H, I)	3.93	6.87	30.12
Store (B, C) / Control (G_2, H, I, J)	3.92	6.84	20.12

* Tabled F-values for the appropriate degrees of freedom.

TABLE IV

MEAN GAINS ON PENCIL-AND-PAPER TEST MADE BY NINTH-GRADE EXPERIMENTAL CLASSES IN THREE SCHOOLS

Schools	Procedures	Gains in SD's
A	Classroom	.96
	Store	1.77
B	Classroom	.43
	Store	.95
C	Classroom	.34
	Store	1.00

Referring again to Table IV it will be observed that the mean gain of the store group was about twice as great as the mean gain of the classroom group in each of these three schools.

high gains had been made and which had been taught by four different teachers. Reference to Table II will show that the store classes from Schools B and D, and the classroom classes from Schools A and F, comprised such groups.

Both classroom and store groups were significantly superior to the control group, as is clear from an analysis of Table V and Table II. Table V shows that the F-ratios of 53.10 and 87.67 were strikingly beyond the .01 levels of 6.90 and 6.84, respectively. Table II shows clearly that the mean achievement of pupils in the experimental classes (both store and classroom) was distinctly superior to that of the pupils in the homogeneous control group, from which School J had been eliminated.

TABLE V

F-RATIOS OF ADJUSTED RETEST SCORES OF TENTH-GRADE EXPERIMENTAL AND CONTROL GROUPS DRAWN FROM NINE SCHOOLS

Groups and Schools	Theoretical F-Ratios (.05)	Theoretical F-Ratios (.01)	Observed F-Ratios
Store and Classroom			
Store (A, B, D, E)	3.91	6.81	6.36
Classroom (A, B, C, F)			
Store (B, D, E)	3.94	6.90	8.99
Classroom (B, C, F)			
Store (B, D)	3.96	6.96	16.53
Classroom (A, F)			
Classroom and Control			
Classroom (A, B, C, F)	3.94	6.90	53.10
Control (G_1, , H, I)			
Store and Control			
Store (A, B, D, E)	3.92	6.84	87.67
Control (G_1, , H, I)			

Tenth Grade.—Table V shows that three different comparisons were made of the achievement of pupils in store and in classroom groups, and that in each comparison the difference was in favor of the groups with store contacts. In the first comparison of the store and classroom groups the F-ratio of 6.36 almost reached the tabled F-ratio of 6.81; and in each of the other two comparisons the observed F-ratio exceeded the .01 level for the tabled F-ratios, being 8.99 and 16.53 in contrast to the tabled F-ratios of 6.90 and 6.96, respectively. In the second comparison, School A was eliminated, because the investigator felt that, in some respects, the experimental conditions existing in School A were not comparable to those existing in School B where the teacher used both store and classroom procedures. In the third comparison were included four classes in which

The gains made by the tenth-grade experimental classes in terms of standard deviations are presented in Table VI.

TABLE VI

MEAN GAINS ON PENCIL-AND-PAPER TEST MADE BY TENTH-GRADE EXPERIMENTAL CLASSES IN SIX SCHOOLS

Schools	Procedures	Gains in SD's
A	Store	2.22
	Classroom	2.52
B	Store	2.62
	Classroom	2.24
C	Classroom	1.44
D	Store	2.88
E	Store	2.23
F	Classroom	2.02

Table VI shows that in both store and classroom groups the gains made on the re-

test over the pretest were greater than 1 standard deviation, and in all but one case they were equal to 2 or more standard deviations. It will be noted that the two highest gains, more than two and a half standard deviations, were made by the pupils who had store contacts in Schools B and D; and that the lowest gains, 1.44 and 2.02 standard deviations, were made by pupils who followed classroom procedures in Schools C and F, respectively.

Supplementary Evidence

Supplementary record forms provided opportunities for ninth-grade pupils in the experimental classes to rate the foods they and their class-mates or teachers purchased, and to check the procedures that they thought should be followed when they encountered certain problems in their shopping; and a month after completion of the foods unit to indicate how much buying they had done in the interim.

Comparative analyses of the records of the ninth-grade store and classroom groups showed that teachers' ratings were consistently similar to those of the pupils, and the mean ratings given on the foods purchased for class use were within approximately three points of the total possible score (21 in one device and 18 in the other). Not only were ratings of pupils and teachers similar, but the mean ratings given on the food purchased in both the store and the classroom groups were very similar. Apparently, the devices were sufficiently simple and objective, so that pupils could rate with them quite accurately, and a satisfactory quality of produce, both staples and fresh fruits and vegetables, was purchased for class use.

Pupils in each class in the store group indicated three times during the experimental unit their judgment regarding five described situations, while pupils in the classroom group did so only at the close of the unit. Therefore, only the ratings made at the close of the unit were used in comparing the judgment of pupils from store and from classroom groups.

There were marked differences in what pupils in the classroom and in the store groups said they did or would do in various situations. For example, when they encountered problems pertaining to price, quality, or kinds of articles on sale, pupils who had had store contacts tended to say that they would judge amounts and qualities and would make substitutions for various types of food items, whereas pupils in the classroom group were more likely to say that they would go to another store to purchase a specific item.

It was not possible for the investigator to determine whether a better selection of food was made when a similar item was substituted or when the specified item was purchased at another store; but their replies suggested that pupils who had had store contacts would be more self-reliant, and more likely to exercise judgment in meeting a variety of situations in the stores, than would the pupils without store contacts.

At the close of the experimental unit, pupils were asked to submit unsigned statements indicating their reactions to the procedures followed when studying selection of food or electrical equipment. The written reactions of pupils showed that those in the store group consistently favored having store experiences as a part of home economics class work. One-third of the pupils from the ninth-grade classroom group stated that going to the stores should have been a part of their class work, and many of them said that they disliked making out grocery orders and felt that this should have been omitted.

About one month after the completion of the unit, the cooperating teachers secured unsigned reports from pupils on the buying they had done during one week. Only about half as many pupils from the store group as from the class group had bought no groceries during the week prior to the date of reporting, and a significantly greater percentage of them made out grocery orders and bought groceries several times during the week. A greater percentage of pupils from the store group also went to the stores to do their buying, although the critical ratio of the difference between the two groups did not reach the five per cent level of significance.

Pupils in the tenth-grade store groups checked themselves after each of the six field trips they took. They used a check list so constructed that the items would be applicable to any piece of electrical equipment, because it was recognized that pupils might investigate a different type of equipment on each trip. Consequently, it was not possible to make comparisons pertaining to similar articles, as was the case in the ninth-grade

in an attempt to evaluate changes that resulted from instruction.

Written reactions of tenth-grade pupils to store and classroom procedures were similar to those of the ninth-grade pupils. The majority felt that they gained a great deal from store experiences, and those in classroom groups believed that they missed a great deal because they had not had field trips to the stores. The majority of tenth-grade pupils in both store and classroom groups listed specific items of information that they had learned, and reported that the booklet, *Selection and Care of Electrical Equipment*, had been particularly helpful.

Conclusions

It seems clear from the findings of this study that community resources can be used in connection with home economics instruction at the secondary level, and that units in consumer buying will be less effective when pupils deal merely with classroom materias than when they have opportunities to make actual contacts with stores, to see the products for sale, and to talk to the merchants or clerks. The differences in the achievement of classes having store contacts and of classes lacking such contacts were consistently in favor of the former. In the ninth-grade classes the differences in mean retest scores were well beyond the .01 level, and surpassed it in two of the three comparisons made of tenth-grade groups and almost reached the .01 level in the third comparison. Supplemental evidence collected from the ninth-grade classes showed that the superiority of the pupils with store contacts was not limited to their achievement on the pencil-and-paper test. Pupils who had had contacts with stores evidenced greater ability to make independent decisions, to form judgments in the light of specific situations, and to carry over into home practice what they had learned in school regarding buying procedures. Apparently the time used in getting to and from the stores was more than compensated for by the values gained from the reality of the situation that visits to stores brought.

Another factor that made for superior achievement was the provision of concise, well-organized instructional materials for pupils to study. This was shown by two facts: (1) the differences between the experimental and the control classes were much less marked in the ninth-grade comparisons than in the tenth-grade (except in the case of School J); (2) the mean gains made during instruction in the ninth-grade experimental classes ranged from approximately one-third to one and three-fourths standard deviations; whereas, the mean gains in the tenth-grade experimental classes ranged from about one and one-half to almost three standard deviations.

The reader will recall that it was only in the tenth-grade experimental classes and in the control class in School J that pupils used especially prepared instructional materials.

The great differences in favor of the tenth-grade experimental groups (both store and classroom) when comparisons were made with the achievement of the control classes (except that in School J) were due in part, also, to the fact that more time was spent on the electrical equipment unit in the experimental than in the control classes. Had the number of hours devoted to the unit in the control and experimental classes been equivalent, the differences in achievement would doubtless have been less marked, but they would undoubtedly still have been in favor of the experimental classes.

Perhaps the final generalization warranted by the findings of the study is that teaching procedures are highly important determiners of pupil achievement and that the effectiveness of school instruction can be materially increased as teachers recognize the value of efficient procedures and learn to utilize them. Up to the present time, too little is known about what teaching procedures will be most useful with different age groups and with different types of material; but it is believed that this study has shown the value at the secondary level of utilizing community resources when they are closely related to the problems with which classroom instruction is dealing.

AN EXPERIMENTAL EVALUATION OF THE SCHOOL EXCURSION[1]

ELLA CALLISTA CLARK
State Teachers College, Winona, Minnesota

INTRODUCTION

Despite the increasing use of the school excursion in the elementary grades, the published literature contains no description of scientific studies designed to evaluate the contributions of the excursion in the elementary school. The scientific viewpoint in education suggests that experimental evaluation be used to test the assumed values commonly attributed to any teaching technique, in order to discover the extent to which the claims postulated for it are realized. This seems to be especially pertinent in the case of the elementary school excursion, which constantly is receiving greater emphasis. There is, of course, a severe limitation of techniques available for a study of this kind, since some results are capable of some objective measurement, while others are extremely difficult to evaluate. Notwithstanding this handicap, undertaking such an experiment appeared both challenging and promising. Accordingly, the present experiment was carried out in an attempt to determine what contributions the excursion might yield in four sixth grade units: Egypt, Printing, Transportation, and Communication. Previous to making the study, the writer conducted a careful search of all published literature relative to the school excursion.

The varying amount of current literature that appears from one period to another may be considered an index of the trend of interest in the school excursion. Therefore, it is interesting to note that up to 1920 only scant reference to the excursion is recorded in the literature. However, the trend from then forward is toward a decided increase in the number of articles appearing. For the five-year period, 1920–1924, only 4 articles appeared; during the interval, 1925–1929, the number rose to 19; while from 1930 to 1934 the articles reached 45; and for the period, 1935–1939, 185 articles appeared. However, despite this rapid increase in the number of published references to the excursion, a perusal of the literature shows that most of the pronouncements concerning the value of the excursion are in the realm of opinion; very few experimental evaluations have been recorded, and none of these represents the elementary school. Yet interest in the excursion has shown unusual growth. In fact, some school systems have established special departments to care for school excursions. Therefore, it appeared that an experimental evaluation of the excursion at the elementary level would make a definite and timely contribution to education.

[1] Report of a study conducted to meet requirements of a Doctor of Philosophy degree, University of Minnesota, 1942.

PLAN OF THE STUDY

The present experiment was carried out in an attempt to identify some of the contributions that the excursion, as used here, yielded in four sixth grade units: Egypt, Printing, Transportation, and Communication. As used in this study, the term, excursion, refers to any carefully planned trip that a group of pupils makes as a part of regular school work, in order to attain specific educational objectives.

In designing the study, statistical treatment of results and the planning of the experiment have been considered as two fundamental aspects of a single problem. Besides, from the outset, a definite effort was put forth to make this a self-contained experiment.

One of the distinguishing characteristics of this study is the fact that the experimental data came from regular classroom situations that were closely planned, but in which a deliberate attempt was made to create no abnormal school situations. It was the purpose of this study to evaluate as well as possible the excursion when used in a natural school environment.

Choice of Units, Locale, and Teachers for the Experiment.—After careful study, as well as preliminary try-out of many different units in various grades, the four sixth grade units of Egypt, Printing, Transportation, and

Communication were chosen, because they seemed to afford promising excursion possibilities and also seemed to represent at least a semblance of comparative difficulty and length. However, the experimental design provided for separate interpretation of the results for each unit. The excursion selected for each of the units was as follows:

Egypt—The Egyptian room at the Minneapolis Institute of Arts.

Printing—A trip through the plant of a Minneapolis daily newspaper.

Transportation—A trip from Minneapolis to St. Paul on the streamlined train, The Hiawatha, and return by bus.

Communication—The main Minneapolis telephone exchange.

Preliminary trial of the experiment in three separate Minnesota schools demonstrated many serious disadvantages, owing chiefly to the difficulty of getting teachers together for meetings, the impracticability of having all schools use the same excursion site, and certain varying administrative features. Inasmuch as experimental design demands that as nearly as possible uniform conditions be maintained for all participating schools, it seemed wise to choose all the cooperating schools from one system. In order to reduce teacher differences and to secure an unbiased evaluation, it seemed advisable to select teachers who not only possessed skill in teaching in the control situation, but who could also use the excursion well. Consideration of all these factors pointed toward the desirability of carrying out the study in the Minneapolis public schools that appeared to afford promising opportunity for maintaining satisfactory experimental conditions. Consequently ten Minneapolis teachers were chosen to participate in the study. Each of these had straight six-A grades. They also evinced sufficient professional interest and alertness to justify the expectation that the extra work involved in the study, planning, and keeping of reliable records requisite to the successful conduct of the experiment would be done in a creditable manner. When the first unit was almost completed, one of the participating teachers had the misfortune of being in a serious accident, which necessitated her absence from school during the remainder of the year; this reduced to nine the number of experimental schools. The pupil distribution in these schools was as listed in Table I. A total of 335 sixth grade children took part in this experiment.

TABLE I

DISTRIBUTION OF PUPILS IN THE EXPERIMENT

School	Boys	Girls	Total
A	16	23	39
B	18	24	42
C	22	14	36
D	19	15	34
E	15	24	39
F	19	22	41
G	12	14	26
H	18	15	33
I	24	21	45
Total	163	172	335

At the same time, another major purpose dominated this study. So often in educational experiments, when several classroom teachers cooperate in a study, they are merely tools for the collection of data that an investigator wishes to secure, and professional growth for the teachers usually represents a minimum. This need not be true. At the same time the teachers' efforts are being used to make a contribution to additional educational knowledge, provision may also be made to give cooperating teachers functional contact with the systematic application of the experimental method of attack on their every-day classroom problems. The growing popularity of the so-called "work-shop" courses attests to the eagerness of teachers to take advantage of opportunities to gain first-hand contact with cooperatively planned experimental evaluation. Accordingly, this study was planned, not only as a device for securing reliable data through collective research, in order to avoid uncritical acceptance of generalizations based on but a single situation, but it was also designed as a teacher-training technique. If this objective were to be realized, two requirements must be met. The necessary theory underlying the experimental method must be carefully organized and presented to the teachers, and the teachers must be given an opportunity to take part in a cooperative development and evaluation of the experimental plan. Therefore, this study was launched as a work-shop application of cooperative evaluation of the excursion technique. A series of meetings provided in-service training of teachers and afforded them an opportunity to study experimental evalua-

tion. It was also the purpose of these meetings to help develop in the cooperating teachers the scientific attitude toward the study, so that they would view in an objective manner this effort to secure experimental proof as to what changes occur in the behavior of pupils who use the excursion, as compared with those who do not.

Measurement of Instructional Effect

No tests were available that would adequately measure achievement in the instructional units of Egypt, Printing, Transportation, and Communication at the sixth grade level. Therefore, after careful setting up of objectives for each unit, tests were constructed and tried out and revised several times in a three-year preliminary study. In order to validate the subject-matter content of each test, each item included was checked for its frequency of inclusion in six sources. The trial form of the test was then carefully edited by several persons in an attempt to evaluate each item on the basis of its functioning content. Following a preliminary try-out of the tests with 100 six-A pupils not included in the final experiment, certain test items were eliminated, if all or none of the children had answered them correctly, or if the discriminating power of any item was inadequate.

To determine the reliability of the achievement tests, the method described by Rulon[2] was used. The formula referred to is:

$$r_w = 1 - \frac{\sigma_d^2}{\sigma_t^2}$$

Every item of the information covered by the test items was presented to both experimental and control groups during the two-week period while the unit was being studied during the final experiment; thus the test favored neither group.

As is shown in Table II, the types of questions used in the achievement tests consisted of completion, multiple choice, matching, and arrangement items. Table II also shows that from 40.7 to 42.5 per cent of the questions in the various tests could have been answered by a child who had no other contact with the materials of instruction than that afforded by the excursion content, which, it must be remembered, was presented to both experimental and control groups.

Tests Other Than Achievement.—As Dewey[3] pointed out:

"Perhaps the greatest of all pedagogical fallacies is the notion that a person learns only the particular thing he is studying at the time. Collateral learning in the way of formation of enduring attitudes, of likes and dislikes, may be and often is much more important than the spelling lesson or the lesson in geography or history that is learned. For these attitudes are fundamentally what count in the future. The most important attitude that can be formed is that of desire to go on learning."

There seems to be general agreement that this is true, and modern education attempts to do something about it despite the fact that measurement in this field is beset with obstacles. Accordingly, in this study it seemed desirable to incorporate into the testing program some means of determining pupil changes of this type. The following four methods were used:

1. For each unit an interest test was constructed; it consisted of from 40 to 60 items, which pupils marked "like", "dislike", or "indifferent". As was the case with the achievement test, the interest test was administered before and after each of the two-

[2] Phillip J. Rulon, "A Simplified Procedure for Determining the Reliability of a Test by Split-Halves," *Harvard Educational Review*, IX (January, 1939), 99–103.

[3] John Dewey, *Experience and Education*, p. 49. New York: Macmillan Co., 1938.

Table II

Analysis of Types of Questions in the Information Tests

Unit	Completion	Multiple Choice	Matching	Arrangement	Total	Excursion Total*	Per Cent of Total Answered by Excursion
Egypt	27 (9)*	30 (14)*	15 (7)*		72	30	41.6
Printing	14 (4)*	27 (9)*	15 (6)*	24 (15)*	80	34	42.5
Transportation	30 (16)*	29 (17)*	15	12 (2)*	86	35	40.7
Communication	25 (8)*	23 (12)*	6 (6)*	18 (4)*	72	30	41.6

* Number of questions the answers to which were supplied by the excursion materials.

week units, as well as a month after the unit had been completed.

2. Pupils' free reactions to each unit were also secured. The day after the culmination of each unit each teacher asked her pupils to write approximately a page giving their reactions to the unit just completed, telling: (1) how they liked the unit, (2) what parts they had found most interesting, and (3) what activities related to the unit they would like to carry on independently.

3. All cooperating teachers kept a diary and anecdotal record of pupils' reactions to the various units as the experiment progressed. From conversation, interview, and observation of the behavior of pupils teachers listed appreciations, leads into further activity, desire to repeat experiences, and any other items which to them seemed pertinent to a complete evaluation of the experiment.

4. Things which the children made during the experiment were saved for later evaluation. Included were objects, models, specimens, poems, stories, lists, and many other interesting products.

Experimental Controls

In so far as possible, the observed differences between the effects of the experimental and control techniques must be due to the experimental factor, if conclusions are to be justified. It follows that maintenance of adequate experimental controls is of fundamental importance to a self-contained experiment. Therefore, a plan was devised whereby it seemed that many non-experimental factors could be fairly well controlled. This was done in the ways described in the succeeding paragraphs.

Teacher Training.—First of all, it was necessary to attempt to insure equally effective teaching in both experimental and control groups; every teacher must understand and carefully follow every step in the procedure which they cooperatively adopted; and the teachers' personal bias either favoring or opposing the excursion must be reduced to a minimum. In an attempt to accomplish these objectives, all teachers had a period of training that preceded, paralleled, and followed the experiment. In these meetings all steps of the procedure were carefully worked out and discussed, the underlying theory with its implications was emphasized, and adequate illustration was given as needed. After each meeting, every teacher received a typed copy of the complete directions in which had been incorporated any worth while suggestions made during the meeting.

Teachers' Diary Record.—Each teacher kept a detailed diary and anecdotal record of how each unit was taught; included also were any pupil reactions that appeared pertinent. In collating and interpreting results, this record proved invaluable, for it facilitated checking on any departure from prescribed procedure that might have influenced the results obtained. It also provided insight into classroom techniques employed, and supplied valuable information concerning pupils' reactions.

Constant Pupil Factors.—For both experimental and control groups the following conditions were maintained throughout the experiment. (1) Both groups had the same content presented to them. Material covered by the experimental group through the excursion was presented to the control group from mimeographed or printed material and with the aid of pictures. In other words, each group was given the best available instruction in each unit. (2) Both control and experimental groups spent the same amount of time (two weeks) on each unit. (3) The same unit was presented in all cooperating schools at the same time.

Rotation of Schools.—In order to reduce further any teacher bias toward the experimental factor, schools were rotated, so that each school which acted as a control group on one unit became an experimental group on the next unit. The rotation plan provided a challenge to participating teachers to do effective teaching in both experimental and control situations, since each teacher taught both types of units. Rotation also tended to reduce the danger that any school which might be superior or inferior could in any way "load the findings".

Zero Control School.—In many studies the question arises as to how much of the demonstrated growth presumably attributable to the experimental factor may be due to such factors as maturation, practice effect, and unreliability of tests. In order to provide a basis for estimating what changes might ordinarily take place in pupils who did not receive systematic instruction in the experimental content, a St. Paul sixth grade consisting of 19 boys and 20 girls was chosen

as a zero control group. The pupils in the zero control group received no instruction in the units concerned; in fact, the teacher made a definite effort to eliminate all such content from her teaching during this period. However, the teacher administered to her pupils all the tests on the same days that the experimental schools took them. It seemed that any growth shown by pupils in the zero control group might thus be attributed to such factors as maturation, practice effect, and unreliability of tests. Hence, in the analysis of experimental results, consideration of this information facilitated a more accurate evaluation of pupil growth due to the instruction involved in the experimental units.

Character of the Experimental Subjects

Intelligence quotients of the subjects of this experiment were secured from the Kuhlmann–Anderson Group Intelligence Test for sixth grade. The mean intelligence quotient for the separate schools varied from 95.73 to 106.41, with a mean of 101.14 and a standard deviation of 9.38 for the total group. Analysis of variance of chronological age indicated homogeneity at the one per cent level. In both intelligence quotient and mental age, analysis of variance indicated heterogeneity. However, in the statistical analysis, mental age is one of the basic matching characters.

As to sex, the experimental subjects were balanced within narrow limits. The slight difference made it appear rather doubtful that it would affect results to any pronounced degree. However, to test that assumption for each unit, a chi-square test of sex distribution was made, and in no unit was a significant sex difference found between the control and experimental groups.

The socio-economic status of each pupil was determined by applying the "Alphabetical Occupational List for Classification of Occupations by Minnesota Occupational Scale"[4] to the data given on the pupil's individual record card. For each unit, application of Brandt's method of the chi-square test indicated that the differences between the distributions of the control and the experimental groups in socio-economic status were not significant.

[4] Florence Goodenough and J. E. Anderson, *Experimental Child Study*, pp. 501–512. New York: Century Co., 1931.

Statistical Techniques

The main statistical techniques employed in this study include chi-square, analysis of variance, the Johnson–Neyman analysis, and the Harris technique. These were used as described in the following paragraphs.

In order to test the significance of the difference in socio-economic distributions of the control and experimental groups in each unit, Brandt's method of chi-square was used. For no unit was a significant chi-square found; therefore, for each of the four units the socio-economic status may be regarded as homogeneous. Another application of the chi-square test was made to determine whether the sex ratio among schools differed significantly. For no unit was a significant chi-square found, thus indicating that for each control group, as well as for each experimental group, there was no significant difference in sex proportion. However, since the requirement of proportionate frequencies was met by the data only approximately, Snedecor's adjustment was used in making the analysis of variance.

Because of the practical difficulty of making each school an experimental unit in itself, it was necessary to determine whether the data from the several control schools could be collated and also whether the same could be done for the experimental schools. In order to make this test of homogeneity, analysis of variance involving two criteria with cross-classification was employed. In cases where heterogeneity prevailed, it was necessary to subdivide groups into homogeneous units. For example, in the unit on Printing, it was found that the experimental schools were heterogeneous according to analysis of variance on final test scores; thus breaking the group into homogeneous units resulted in having two experimental units, each of which was then compared with the control schools.

However, before the method of analysis of variance could be applied, it was necessary to test whether the variances within groups were uniform. To test whether this condition prevails when only two groups are under consideration, the exact statistical test is the Z-test. When more than two groups are under comparison, the approximate test suggested by Welch with the use of Nayer's tables of the sampling distribution of the criterion L_1 can be employed.

The Johnson–Neyman technique is the basic statistical tool used in interpreting the data obtained in this study. The functions of this technique are: (1) to ascertain whether the experimental group differed significantly from the control group on the final test scores and (2) to determine the region of significance which defines that part of the population for which any differences found are significant. The Johnson–Neyman technique also permits use of all data collected.

It will be recalled that the analysis of variance indicated heterogeneity for the schools as a whole on mental age; therefore, there appeared need for controlling this characteristic. High correlation was found to exist between pretest and final test; for this reason it seemed advisable to match on pretest scores as the second basic characteristic. In this statistical analysis, then, the basic matching characters are mental age and pretest score, and the final score (obtained at the end of the two-week unit) is the criterion.

As previously stated, four weeks after each unit was finished, the achievement test was given again to measure retention of information at the end of that interval. In order to determine whether pupils who had the highest final scores (at the end of the unit) also had the highest retention scores, the product-moment coefficient of correlation was computed for both experimental and control groups. The degree of relationship shown by the correlation coefficient indicated the extent to which those pupils who had the most information at the end of the unit remembered absolutely more at the end of the four-week interval. However, these coefficients do not answer the question of whether the individual who had the most information at the time of completing the unit retained relatively the most four weeks later. To obtain this information the Harris technique[5] was used.

Findings of the Experiment

Achievement.—In the Egypt unit no sex differences appeared. The control group was significantly superior to the experimental group on the basis of final test scores when mental age and pretest scores were held constant. The region of significance is bounded by an hyperbole. Within the region of significance the value of F is favorable to the control group. That value is significant for that part of the population which falls within the hyperbole. Here the level of significance is F = .01. In other words, pupils whose initial scores and mental ages lie within the region benefitted significantly more if they were in the control group than if they were in the experimental group. Broadly speaking, pupils in the range from 0 to 37 on pretest and from a mental age of approximately 110 to 167 profited more from the control method. The best estimate of the true difference between the means of the control and experimental groups is 14. There was no significant difference in either absolute or relative retention for the experimental and control groups.

In the Printing unit no sex differences were found. For the control group, analysis of variance of final test scores with two criteria involving cross-classification indicated homogeneity; therefore, the control schools could be collated and regarded as homogeneous. However, in the experimental group the schools were not homogeneous (F = 5.5, which exceeds even the one per cent tabled value). Thus, it was necessary to subdivide the experimental schools into groups that were homogeneous. The t-test indicated that the experimental schools could be subdivided into two homogeneous groups. The Johnson–Neyman analysis was then used to compare the final test results secured in each of these two experimental groups with the final scores secured in the control group. In both cases the means of final scores for the two experimental groups were significantly the greater when adjusted for pretest scores and mental age. For the first experimental group a region of significance bounded by an ellipse exists at the one per cent level. Those persons whose mental ages (100–140) and pretest scores (0–37) fall within the region of significance profited significantly more if they were in the experimental group than if they were in the control group. The best estimate of the true difference between the means of the control group and the first experimental group is 4.

In the case of the second experimental group in the Printing unit a region of significance (an ellipse) exists at the one per cent level favoring the experimental group. Broadly speaking, persons having mental ages of from 130 to 137 and pretest scores of

[5] Arthur J. Harris, "The Correlation between a Variable and the Deviation of a Dependent Variable from Its Probable Value", *Biometrika*, VI, 438–443.

from 4 to 44 may be expected to do better on the final measurement, if they are members of the experimental group. The best estimate of the true difference between the means of the control group and the second experimental group in the Printing unit is 9.

So far as retention is concerned, in the comparison of the control group with the first experimental group, the control group was superior in absolute retention; while, in the comparison of the control group with the second experimental group, there was no significant difference. In relative retention there was no significant difference in the Printing unit.

The Transportation unit was the only one of the four units that revealed sex differences; these appeared in both control and experimental groups, thereby necessitating analysis separately by sex. The Johnson–Neyman analysis of final test scores of the control boys and the experimental boys, using mental age and pretest score as basic matching characters, shows results favorable to the experimental boys. A region of significance that is bounded by an hyperbole exists at the five per cent level. It indicates that, in general, boys with mental ages of from about 110 to 125 and pretest scores of from about 35 to 60 would tend to profit more from the experimental method than from the control method. The best estimate of the true difference between the means of the control and experimental boys is 1.9. In relative and in absolute retention there was no significant difference between the scores of the boys in the control and experimental groups.

Similarly, a Johnson–Neyman analysis was made of girls' scores in the control and experimental groups for the Transportation unit, but there was not sufficient evidence to reject the hypothesis that there is no real difference between the means of final scores for the two groups when of the same initial scores and mental ages. In absolute and relative retention there was no significant difference between control and experimental groups of girls.

In the Communication unit the Johnson–Neyman analysis of final test scores for experimental and control groups showed that the experimental method was superior to the control method. A region of significance bounded by an ellipse exists at the one per cent level. In general, all pupils with mental ages of from 131 to 174 and pretest scores of from 17 to 54 may be expected to profit significanty more if in the experimental rather than the control group. The best estimate of the true difference between the means of the control and experimental groups is 5.9.

In the Communication unit there was no significant difference between the control and experimental groups in either absolute or relative retention.

Zero Control School.—As explained above, a St. Paul sixth grade had been selected to act as a zero control group. These pupils had no contact with the experimental instruction, but they took all the tests at the same time the children engaged in the experiment took them. For each of the four units, data concerning tests of significance of the difference between the means of the several tests on each unit for the zero control group indicate that in no case was a difference found to be significant. Therefore, it is highly probable that, in this experiment, obtained gains were due to instruction rather than to outside influences.

Interest Tests.—For each unit the responses of each child on the pretest, final, and retention tests were tabulated and the number of responses of "like", "indifferent", and "dislike" for each item by every child was summated. After the same procedure was carried out for both experimental and control groups, a chi-square test was applied. In no unit were the differences obtained significant.

Children's Written Reactions.—It will be recalled that at the end of each unit each of the participating pupils wrote approximately a page: (1) telling how he liked the unit, (2) mentioning the items that had proved most interesting, and (3) listing activities he wished to carry on independently now that the unit was finished. An attempt was made to secure the children's free reactions; teachers were asked to make no suggestions as to what pupils should include. Every child was requested to write exactly as he felt about it. For all units the general tone of the written reactions reflects sincerity and considerable thoughtful evaluation. In order to interpret these data, an attempt was made to analyze the pupils' statements objectively. For example, topics that pupils listed as most interesting in the unit were arranged in order of frequency in a table, which also indicated the number and percentage of boys, girls, and

total pupil group in each of the experimental and control groups that mentioned the item. Since children's written reactions appeared to be particularly interesting for every unit, a resumé of their responses, as made at the end of the Printing unit, is given here as an example. There is no particular reason for choosing the Printing unit, for all four units yielded results of similarly high interest, so any one of them may be regarded as typical.

For the Printing unit the excursion group listed 22 different topics as being most interesting; only eight of these were mentioned by the control group. On the other hand, the control group listed 17 topics, nine of which were not included in the 22 given by the excursion group. In the experimental group 54 per cent of the pupils stated that the excursion was the most interesting part of the Printing unit. Furthermore, having seen the machines in operation, it is not surprising that 11 of the topics the experimental group listed concern machines. On the other hand, the control group, having viewed pictures while they heard an explanation of the operation of these machines, failed to include them as frequently among the most interesting items. Here, as in the other units, pupils participating in this experiment tended to express greater interest in objects they had actually seen than in those presented to them verbally and pictorially. Apparently seeing actual objects makes these objects more meaningful to the pupils than if a vicarious experience is substituted. Pupils' comments as quoted later further emphasize this fact.

In addition to topics that pupils listed as being most interesting, the activities they evinced interest in carrying on independently after the unit was finished are also indicative of the value they attach to the unit. The extent to which any school experience affects pupils' subsequent activities is to a certain extent a measure of its richness. As was done in the case of topics the pupils listed as most interesting, a similar frequency table was made, showing the activities they expressed a desire to carry on independently. Again using the Printing unit as an illustration, it is interesting to note that, although 54 per cent of the experimental group had listed the excursion as the most interesting item in the Printing unit, 21 per cent stated that they wished to revisit the newspaper plant. In this desire they are consistent in attaching value to the excursion. Likewise, 24 per cent of the control group expressed their desire to visit a newspaper plant, indicating a felt need for direct experience such as the excursion offers. Children in the experimental group mentioned with greater frequency that the Printing unit had proved interesting. Careful consideration of all factors involved in the unit, as it was taught in this study, reveals no non-experimental factor to which this difference in attitude may be attributed; therefore, it appears probable that the experimental children liked the unit better than did the control pupils, because they had taken the excursion. Teachers' diaries and anecdotal records connected with the unit further emphasized this probability.

It is difficult to convey an accurate impression of how pupils seemed to feel about a unit, unless one quotes them directly. Acccordingly, a few typical excerpts from pupil's written reactions to the Printing unit are quoted below. The assignment that preceded the writing of these papers, it should be recalled, included no mention of the excursion. Pupils who referred to it did so voluntarily. They volunteered such statements as the following:

"I enjoyed the excursion to the newspaper plant immensely and feel that our whole class learned much more from going to the plant than we would have understood in any other way."

"It is very interesting to watch the bundles of papers being tied with wire. It is clever the way it works. I like to watch the matrix being made. It's all very interesting. I hope I can go again sometime."

"In the study of Printing I liked the trip very much. It was fun to go to a real newspaper office to see a real newspaper printed."

"Visiting the newspaper plant increased my vocabulary a great deal."

"I have enjoyed finding out new things about printing more now than before the excursion."

Several pupils in the control group referred to vocabulary difficulties encountered, and some suggested a desire to secure meaningful impressions of these new words by visiting the newspaper plant. They said:

"Some parts of the Printing unit I liked; others I didn't. There were so many words

I can't remember. I can't remember the difference between a chase and a cut. I wish we could really visit a newspaper plant."

"I liked the unit pretty well, but it was very hard to keep the words and their meanings straight. Some day I mean to go to a newspaper plant and see how the daily newspaper is made."

"I hope my father takes me to see how a newspaper is printed. It must be fun to see all the processes right before your eyes."

Indicative of an appreciation of certain aspects of printing, the following statements are quoted from reactions of pupils who took the excursion to the newspaper plant:

"The study of printing was extremely interesting. The wire-photo, teletype, engraving, and linotype were the things that particularly interested me during the trip through the newspaper office, and it isn't until you see it before your own eyes that you realize how much work is put in on the few pages of paper that you can sit and enjoy at home at such little cost. It is not until you have the facts before your eyes that you realize how hard the men employed by the newspapers have to work to be in the position where they are today. From the editors even down to the men who run the presses they all work hard and do their best to publish the best newspaper possible. These men should be congratulated on the splendid example they have set for American industry."

"I like Printing better than the Egypt unit. Since I visited the newspaper plant, I know how hard it is to print books and newspapers. From now on I am going to take better care of them."

"When I got home from the newspaper plant and read the paper, it seemed more interesting to know how many people work to give us the news we read."

"Since my visit to the newspaper plant, I take more interest in the newspaper; I don't just read the funnies any more."

"Since I have been through a newspaper plant where I saw all the work they put on the paper, it makes me much more interested in the paper. I read many more parts of the paper now because I think it is more interesting after you see how much work they put on it."

Summary of Conclusions

The purpose of this study was to determine what contributions the excursion might yield in four sixth grade units: Egypt, Printing, Transportation, and Communication. In no unit except Transportation were sex differences found. In a comparison of control and experimental groups, in means of final scores when adjusted for mental ages and pretest scores, the following groups were found to be significantly superior:

Egypt—control group.
Printing—experimental groups.
Transportation—experimental boys.
Communication—experimental group.

There were no significant differences between control and experimental groups in absolute retention, except in the case of the Printing unit in which the control group was significantly superior to one of the two experimental groups.

In relative retention there was no significant difference between control and experimental groups in any unit.

Results secured with a zero control group that had no classroom instruction in the content of any unit, but took all tests on schedule, indicated that factors such as maturation and practice effect had not influenced significantly test scores for these units. Therefore, it is highly probable that in this experiment obtained gains were real gains due to instruction rather than to outside influences.

Chi-square tests of the distributions of responses of "like", "indifferent", and "dislike" for items in each interest test revealed no significant differences between the control and experimental groups.

Children's voluntary reactions written at the end of each unit indicate that the pupils who went on the excursion recorded for each unit a greater variety of items which especially interested them. They also listed with greater frequency a wider variety of activities in which they stated they would like to engage after the unit was finished.

Educational Implications of This Study

The evidence presented through this study points to the definite need for careful preparation in the use of the excursion. In this investigation not only were teachers and pupils carefully prepared to secure maximum benefits from each excursion, but at the same

time the administrative aspects were cared for, perhaps as effectively as any similar well-organized set-up could provide. It is pertinent to remember that the gains found in this study were achieved under conditions in which: (1) an appropriate excursion was carefully selected; (2) teachers had received systematic in-service training in the use of the excursion and its underlying psychological implications; (3) pupils had developed a purposeful attitude toward the excursion as valuable source material; (4) administrative difficulties had been reduced to a minimum; (5) appropriate guide service was provided on each excursion; and (6) the results of the excursion were utilized in subsequent class discussion. This study supplies no basis for the assumption that positive results will be obtained with excursions, unless similar conditions of careful planning are fulfilled.

On the other hand, it must be borne in mind that the control groups in this study also received the best possible instruction under the given circumstances. The control group had contact with all the information that the experimental group saw on the excursion. In all probability, few teachers not using the excursion would have this material available.

The findings of this study further suggest that teachers must constantly evaluate the excursions they use. Some of them may not be bringing about the hoped-for results. In an elementary-school survey, Borgeson[6] found that excursions to museums were the most popular school excursions. Supporting this practice, Atyeo's study[7] found further that, in a museum type of unit, results definitely favored the experimental group of ninth grade children. By way of contrast, in the present study, the museum type of excursion was the only one of four excursions studied in which differences favored the control group. Together these studies constitute an argument against generalizing for all situations on the basis of either present practice or the evidence presented by one experiment.

Perhaps one of the paramount values of the excursion, aside from its success in aiding in the acquisition of knowledge, is another aspect which does not lend itself readily to evaluation; that is, its power to develop interests, ideals, and appreciations. The findings of this study indicate that the pupils who took the excursion showed at the end of the unit evidence of broader interest than pupils who did not take the excursion. Of course, this furnishes no evidence of the extent to which these declared interests eventuated in worthwhile subsequent activities. That remains a problem for future investigation in a field which bristles with interesting challenges.

Another implication that may be of particular interest to the experimental research worker emanates from teachers' reactions to having cooperated in such a study on the work-shop basis. All teachers concerned expressed definitely favorable attitudes toward the opportunity of having been participants in a study in which the emphasis was placed not merely on the collection of factual material for someone else's use, but also upon the round-table method of developing and carrying out a scientific study. All teachers stated that they had expended an enormous amount of time and energy on the work, but they also stated that the experience had been a decided challenge to them. The teachers felt that through participation in the study they had developed valuable insights into the scientific method of evaluating and improving their own teaching. Several requested an opportunity to secure further in-service training in this same way.

Anyone who attempts to do a piece of research involving children should be definitely concerned with the effects upon the child. Experimental research is necessary to the evaluation of educational procedure, but the subjects to be used in the experiment should also receive due consideration. In the present study, it seemed that, because of careful planning, pupils of both control and experimental groups were at all times spending their time to advantage. Children's written reactions to the units included many expressions of favorable attitudes toward the procedures being used, with no adverse criticisms. Therefore, it seems that, with serious consideration of the probable effects which a piece of experimental research may have upon children, an experimenter can do much to reduce the undesirable effects that otherwise might be concomitants of his study.

[6] F. C. Borgeson, "Excursions in School Life," *Bulletin of the Department of Elementary School Principals*, XI (April, 1932), 461–66.

[7] H. C. Atyeo, *The Excursion as a Teaching Technique*. Teachers College Contributions to Education, No. 761. New York: Teachers College, Columbia University, 1939. Pp. 225.

COMPARATIVE EFFECTIVENESS OF PICTORIAL TEACHING AIDS

An Experimental Investigation in Safety Education at the Elementary-School Level*

DAVID J. GOODMAN
Educational Specialist, Training Aids Division
Army Air Forces, New York City

ANALYSIS OF PROBLEM

Statement of the Problem.—The purpose of this study was to determine the comparative effectiveness of certain diverse forms of pictorial teaching aids when used in teaching certain topics in safety educaton to pupils of the sixth grade of the elementary school.

The pictorial teaching aids used in this study are all devices for group instruction using projection equipment. They are the sound motion picture, silent motion picture, silent film slide, and sound film slide. The four subjects on safety selected for use in the investigation were: automobile safety, fire safety, pedestrian safety, and bicycle safety.

Specific Problems.—The major problem divided itself into specific problems:

1. How do the different visual media compare in effectiveness in the presentation of factual information on the several safety subjects? Is there any real difference between them to justify the setting up of special criteria for their use?

2. Can the four safety subjects be comprehended and mastered equally well by pupils through the use of the visual media employed in this study, or do any of the subjects lend themselves particularly well to study through the use of the pictorial aid?

The problem was, therefore, to determine differences in the instructional outcomes for the media used, and the degree to which the differences are significant, rather than to determine whether pictorial aids are effective in teaching.

Significance of the Study.—The value of pictorial teaching materials as aids to learning is no longer contested. Not only educators but practically all government branches, both military and civilian, and business organizations have recognized their effectiveness in an educational program.

Since teachers and others are using the various pictorial media, in most cases indiscriminately, it is desirable that they have a basis of reliable research upon which to determine the most suitable medium for use in a given situation, to discover whether real differences do exist among the media and, if so, where those differences lie.

Safety was the subject selected in this investigation, since it is one which, in recent years, has received much pictorial presentation. This is understandable. Safety concerns itself primarily with the actions of people. It is essentially a "picture" study and, therefore, lends itself particularly well to instruction through the use of pictorial materials. In addition, since the subject deals with accidents, involving life or death situations, special attention has been given to instruction in safety for children, who need a fundamental understanding and knowledge of safe habits and a strong "conditioning" for the development of proper safety attitudes.

Some research evidence exists regarding the comparative effectiveness of two or three pictorial media; e.g., the motion picture and the still picture. However, in these studies, the pictorial material appearing in the two media was seldom the same or even equivalent; that is, the still pictures were not "lifted" from the film. Yet no *real* comparison can be made unless identical or equivalent material appears in the media being compared. In the present investigation the media have been made equivalent in terms of pictorial and textual content. The investigation attempts also to throw much needed light on the effectiveness of the "new" media, sound film slide and silent film side, as tools of instruction.

* Abstract of thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the School of Education of New York University, 1942.

Previous Research

A study of previous research in the field of visual education fails to reveal conclusive evidence that any one form of visual aid is more effective than another. The several studies reported by David Sumstine,[1] Joseph Weber,[2] Frank N. Freeman,[3] Daniel C. Knowlton and Warren Tilton,[4] Philip Rulon,[5] Clarence Clark,[6] Varney Arnspiger,[7] Nathan Doscher,[8] and Katherine Rose[9] do reveal that visual media make significant contributions to the learning process; but research evidence showing conclusively the comparative effectiveness of the sound motion picture, silent motion picture, sound film slide, or silent film slide is practically nonexistent. It should be noted that in none of the experiments listed above is the silent or sound film slide dealt with, nor was the experimenter able to locate any studies in which the comparative effectiveness of these devices was investigated.

Description of the Experiment

In organizing and setting up the investigation, one problem above all others required special consideration; namely, that the four different visual aids comprising each set of materials on a specific safety subject contain equivalent factual information. Since a comparison was made between the different media within a set, it was essential that the sound motion picture, silent motion picture, silent film slide, and sound film slide which comprised the set of media on each subject contain the same photographic and textual data, but differ only in the physical form and characteristics inherent in each medium.

One such set of media was prepared for each of four different subjects, as follows: (a) auto safety, (b) bicycle safety, (c) pedestrian safety, and (d) fire safety.

[1] "Comparative Effectiveness of Visual Education in the High School," *School and Society*, VII (February 23, 1918), 235–238.

[2] *Comparative Effects of Some Visual Aids in Seventh Grade Instruction*. Chicago: Educational Screen, 1930.

[3] *Visual Education*. Chicago: University of Chicago Press, 1924.

[4] *Motion Pictures in History Teaching*. New Haven: Yale University Press, 1929.

[5] *The Sound Picture in Science Teaching*. Cambridge: Harvard University Press, 1933.

[6] "Sound Motion Pictures as an Aid in Classroom Teaching." Unpublished Ph.D. thesis, New York University, 1942.

[7] *Measuring the Effectiveness of Sound Pictures as Teaching Aids*. New York: Bureau of Publications, Teachers College, Columbia University, 1933.

[8] "A Critical Analysis of Some Visual Aids Used in Teaching Pedestrian Safety." Unpublished Ph.D. thesis, New York University, 1939.

[9] "Motion Pictures vs. Lantern Slides in Teaching Certain Units of Biology," *Educational Screen*, XXI (January, 1942), 30–31.

The basic medium used to determine the organization and development of the other materials was the sound motion picture, since it was possible to produce the other media working from a sound film, but impossible to prepare a sound motion picture from the other media used in the experiment. The sound film, therefore, had to meet a number of physical and educational requirements.

After much search and careful analysis, four different films were obtained, each on a different subject, each one reel in length, each for approximately the same age level, each available in 16mm. and 35mm. size, and with the 16mm. size of each available in both sound and silent versions, so that the materials of all the media might be equivalent.

To prepare the film strips, the investigator selected from each sequence in the 35mm. motion picture the specific frame which in his judgment most effectively illustrated the safety concept being prescribed. In this process approximately twenty-five photographic frames were lifted out which presented the safety story of the film. Approximately twenty textual frames were then selected from the film which related specifically to the photographic frames previously selected.

Each of the individual frames, photographic and textual, was framed, mounted, and bound between two-inch square cover glasses, producing a set of two-inch square lantern slides. The "textual" slides were inserted in their proper sequence among the photographic slides, and this lantern slide set was a "still" copy of the motion picture in which only the essential photographic and textual frames were used. In essence the lantern slide set was a continuous story exactly comparable with a silent film strip, except that it was not on a single strip of film. The same procedure was followed in preparing the four different lantern slide sets (or silent film slides).

The sound slide film was prepared by recording all the textual frames of each slide set on a special disc which was synchronized with the photographic frames of the film slide set. Picture for picture and word for word, one written and the other spoken, the two film slide media were identical.

A committee of three experts viewed the materials to determine whether the most essential safety concepts of each film had

been selected for use in the slide set, and whether the frames selected on the basis of safety content were also the most effective visually in each film sequence used. All their suggestions were incorporated in the final set of material. The investigator next prepared four separate tests, one for each of the safety subjects selected. Each test consisted of fifty multiple-choice questions, selected after an analysis of the contents of each frame used in the lantern slide set. The fifty questions of each test varied in degree of difficulty. The tests were validated by administering them to six classes of sixth-grade pupils of New York City schools. The questions were also checked by sixth-grade teachers for suitability of vocabulary and sentence structure for that grade level. The easiest questions were then eliminated, so that the final tests comprised 35 questions. Two different forms of each test were prepared. Form A was to be used in the pre-test part of the actual experiment. Form B was to be administered in the post-test (immediate recall) and thirty-day (delayed recall) portions of the experiment. Both forms of each test consisted of the same 35 multiple-choice questions with three possible answers for each question, but differed only in the order of the questions and in the position of the one correct answer for each item.

Eight 6A classes in four elementary schools within adjacent districts in the Borough of Brooklyn in New York City, numbering 240 pupils in all four schools, participated in the experiment. Approximately 130 of the pupils were from classes of high I.Q., while the remaining 110 pupils were from low I.Q. classes. Each school had one class of low I.Q. and one class of high I.Q. pupils. The range and mean I.Q. scores of all classes within each I.Q. group were fairly comparable.

The investigator prepared a special schedule allocating the showing of each set of media. It was planned so that each class in the schools was exposed to a different visual medium on a different safety subject for each of four consecutive days. There was no repetition either of the medium or the safety subject; that is, no class saw more than once any one type of medium or more than one type of medium on one subject. This technique of experimental research is known as the rotation-group method. It has the advantage of eliminating a number of variable factors, such as differences in number of pupils, age, I.Q., etc.

The regular classroom teacher did not enter into any phase of the experiment. The investigator himself conducted the experiment in each of the classes, going from school to school. Thus the factor of teacher contribution in the use of visual media in a lesson was eliminated from the experiment.

One regular class period of forty-five minutes was devoted to the presentation of each medium. The period was broken up into three parts: part 1, first fifteen minutes, pupils answered pre-test; part 2, second fifteen minutes, presentation of medium (no pupil or teacher discussion); part 3, third fifteen minutes, pupils answered post-test.

Thirty days after presentation of the media, again without any teacher or pupil preparation, the subjects in all the classes answered the post-test a second time.

Statistical Treatment and Analysis of Data

Although 240 pupils took part in the study, only 144 of them answered every one of the tests. Only these 144 sets of scores were used for calculating all results and as the basis of all conclusions. The statistical treatment divided the data into two special groups—those pertaining to study of the different media and those pertaining to study of the different subjects. Within each group there was a further breakdown in terms of all groups, groups of high I.Q. and groups of low I.Q. Correlation tables were made of all data, since all scores were correlated. The statistical treatment deals with two specific problems. The first problem was to determine whether the gains registered between two sets of scores—the pre- and post-tests, and the pre- and 30-day test scores—were significant or reliable and, second, to determine whether there existed any differences between the media or any superiority of one medium over another in gains registered by two groups exposed to two different media. The statistical techniques used in arriving at the conclusions of the first problem were mean scores. For each set of data, standard deviations, coefficient of correlation, standard error of difference between two means, actual difference between two means, and reliability of difference between two means (critical ratio) were found. Whenever this ratio is 2.7 or

larger, it means that the difference between the two groups compared is greater than zero, that it is significant. With respect to the second problem, differences between media, the following statistical techniques were used: mean gains in a set of scores, standard deviations of the mean gains, the standard error of the differences between the two mean gains, and finally the critical ratio of the two mean gains. Again, if the critical ratio was 2.7 or more, it may be assumed that the gain was a reliable one and that the medium registering this gain could be considered superior to the one with which it was compared.

Conclusions

Effectiveness of the Media—All Groups.—1. All four media are effective devices for presenting safety information, as evidenced by the high critical ratios obtained, indicating that real and highly significant gains were produced for both immediate and delayed recall.

2. The silent motion picture produced the largest gains in test scores, for both immediate and delayed recall.

3. The silent and sound film slide followed the silent motion picture in size of gain produced, on both immediate and delayed recall. There was practically no difference between the two media in this respect.

4. The sound motion picture produced the least gain in test scores in both instances.

5. The silent motion pictured showed a statistically reliable amount of superiority over the sound motion picture and sound film slide for immediate recall, and the indication was less sharp but still strong that it was superior to both for delayed recall.

6. There was a strong indication also that the silent motion picture was more effective than the silent film slide for delayed recall, although not for immediate recall.

7. The silent film slide was clearly more effective than the sound media for immediate recall, but there was no evidence that it was superior to them for delayed recall.

8. There was no clear indication that either one of the two sound media was more effective than the other, for either immediate or delayed recall.

High I.Q. Group.—1. All four media were effective in the presentation of safety information for both immediate and delayed recall. Although there was some loss over the thirty-day period, of the initial gains produced, nevertheless, all gains for both immediate and delayed recall were real and highly significant, as shown by the high critical ratios obtained.

2. The silent film slide produced the largest gain in test scores, for both immediate and delayed recall.

3. The sound film slide ranked next for immediate recall, followed by the silent motion picture, but for delayed recall this order was reversed.

4. For delayed recall, the difference between the two silent media was slight, as was that between the two sound media.

5. The sound motion picture produced the largest gain in test scores, for both immediate and delayed recall.

6. The silent motion picture showed a statistically reliable amount of superiority over both sound media, for immediate and for delayed recall.

7. There was a strong but not absolutely reliable indication that the silent motion picture was more effective than the silent film slide for delayed recall, but not for immediate recall.

8. The silent film slide was clearly more effective than the two sound media for immediate recall, but the indications that it was superior to them for delayed recall, while fair, were not wholly reliable statistically.

9. There was no clear evidence that either one of the sound media was superior to the other, in either instance.

Low I.Q. Group.—1. As was shown above for the other groups, here also all four media proved effective devices for the presentation of safety information, for both immediate and delayed recall.

2. For immediate recall, the largest gain in test scores was produced by the silent motion picture, followed by the sound film slide and closely behind that by the silent film slide.

3. The least gain for immediate recall was produced by the sound motion picture.

4. For delayed recall, the picture was sharply altered from that given by the preceding I.Q. groups. For the first time at any point in the data, the sound media recorded the production of a greater gain. The largest amount of gain was produced by the sound film slide, followed by the silent motion picture and, almost equal with that, the sound motion picture.

5. No one medium showed any degree of superiority to the others that was statistically reliable.

6. For immediate recall, the silent motion picture gave some indication of superiority to the two sound media, and somewhat less so to the silent film slide, but none of these indications approached statistical significance.

7. No reliable indications at all were obtained that there was any superiority of one to another, as between the silent film slide and the two sound media, or of the sound media with each other.

8. For delayed recall, the findings were completely reversed. There was some indication that the sound media may have greater effectiveness than the silent, particularly than the silent film slide. But none of these results approached closely to statistcal significance.

Suitability of the Subject Matter—All Groups.—In presenting these conclusions, it must be noted that for each subject only *one* film was used, from which the other three forms of the media were prepared. Therefore, the following conclusions apply only to *these four films and no others.* No claim is made that the use of other materials on these subjects will produce the same results, and no general conclusions on comparative suitability of the different subjects will be made. But, since the materials of the four sets of media were considered fairly equivalent to each other, it was worth while considering whether any indications emerge that certain subjects are superior to others in suitability for visual presentation.

1. All four subjects—automobile, bicycle, fire, and pedestrian safety—registered an effective gain when presented through visual media.

2. There was some loss of information learned on each subject, over a thirty-day period, but all gains made on all subjects for both immediate and delayed recall were significant, as shown by the generally high critical ratios. The subjects on which the greatest amount of gain in mean scores was produced, for both immediate and delayed recall, were fire and bicycle safety.

3. The subject on which the least gain was produced for both immediate and delayed recall was pedestrian safety.

4. The one subject which showed a clearly smaller degree of suitability for visual presentation than the other three was pedestrian safety, both on immediate and on delayed recall.

5. There were some indications that fire safety and particularly bicycle safety were superior to automobile safety for immediate recall, but these differences were not entirely reliable statistically.

6. For delayed recall, no such indications were evident at all.

7. No reliable indications of any superiority as between bicycle and fire safety were given for either immediate or delayed recall.

High I.Q. Group.—1. As with the "all" grouping, all four subjects showed effective gains when presented through visual media.

2. All gains recorded on all subjects were highly significant.

3. On immediate recall, the largest gain was produced on bicycle safety followed by fire safety. However, all critical ratios were very high with small differences among them.

4. For delayed recall, bicycle and fire safety were practically identical in size of gain produced, but greatly exceeded the other two subjects.

5. On immediate recall, no clearly reliable indications of the superiority of one subject over another were found. Some indications did appear that pedestrian safety is less suitable than the other three subjects.

6. On delayed recall, there was a clear, reliable indication that pedestrian safety was less suitable than fire and bicycle safety, and an indication that it may be less suitable than automobile safety.

7. Both fire and bicycle safety gave very strong though not absolutely reliable evidence of superiority to automobile safety, but no indication of any difference between themselves.

Low I.Q. Group.—1. For this group, as for the other two, presentation by means of visual media produced an effective gain for all four subjects.

2. All gains made, for both immediate and delayed recall, were significant, except that for pedestrian safety on delayed recall.

3. The subjects on which the greatest gains were produced, for both immediate and delayed recall, were fire an bicycle safety.

4. The least gain was produced in both instances on pedestrian safety.

5. The one subject which showed a clearly smaller degree of suitability for visual pre-

sentation than the other three, for immediate recall, was pedestrian safety. On delayed recall, it was also clearly inferior to fire and bicycle safety, but the indications were not as definite of its inferiority to automobile safety.

6. There were some indications that fire and bicycle safety were more suitable than auto safety on delayed recall, but the data were not clearly reliable statistically.

7. There were no indications of any superiority as between fire and bicycle safety on either immediate or delayed recall.

Generalizations.—A number of important generalizations, suggestions, and specific criteria emerge for the use of these media in safety education.

1. The investigation offers additional research evidence that all the pictorial media aid materially in the learning process, for both immediate and delayed recall, and for groups of both high and low I.Q.

2. The two comparatively "new" media, the sound and silent film slide, are effective visual tools that should receive the recognition and the use which they merit. They compare favorably with the silent motion picture, and in all instances but one in this experiment were superior to the sound motion picture.

3. The sound motion picture is the least effective of the media for achieving the type of result in safety education used as a measure in this experiment. Perhaps, because of the complexity of the device, the greater requirements for its preparation and use, and the prestige and example of the commercial entertainment film, it would appear that the sound motion picture has had an undue amount of emphasis placed upon it, and its value in the classroom has been too readily assumed. This is an important subject, and a careful measure of its effectiveness should be made, when criteria are used other than the increase in the pupil's fund of information.

4. The silent motion picture is the single most effective medium. The silent and sound film slide was very close behind. Therefore, it is advisable to consider certain factors other than effectiveness in presenting information that enter into the selection and use of a particular visual medium in the classroom.

Costs of projector and film subjects are considerably higher for the motion-picture media than for the film-slide media. The film-slide projector is a much simpler machine, less bulky, less mechanically imposing, easier to operate, and therefore more readily accessible and consuming much less of the brief time of a classroom period than the motion-picture projector.

These points serve to emphasize the advisability of giving more attention and recognition to the film strip as a teaching tool, but they are not set forth to disparage in any way the merits of the silent motion picture. Its effectiveness is clearly established in this experiment. To the teacher or school official, however, these considerations of cost and use are highly important. They must determine whether the difference in learning produced warrants the use of the motion picture in spite of these other factors.

CRITERIA FOR EVALUATING READING TEXTBOOKS

DAVID KOPEL*
Army of the United States
and
JOHN F. O'CONNOR
Army of the United States

During the last decade teachers, administrators, and research workers in the United States have begun to recognize that junior and senior high schools share with elementary schools a definite responsibility for the improvement of reading. This realization has resulted in the development of courses in reading in many American secondary schools.

Conceived first as emergency measures to help "poor readers" attain the reading skills needed to cope successfully with a high-school curriculum, these courses have been broadened in scope as their directors learned that the difficulties and needs of the "poor reader" differed only in degree from those of other children. The first courses were designated "remedial" and were devoted to the treatment of "reading disabilities". Today in some schools provision is made in one or several courses for corrective and developmental, as well as remedial, reading. "Remedial" reading refers to the diagnostic and instructional procedures used with children seriously retarded in reading. "Corrective" reading is designed to aid children who display minor reading deficiencies. For all other pupils, whose present status is satisfactory, a program to promote growth in reading skills and interests — i.e., developmental reading — is provided.[1]

The introduction of reading courses has led to the preparation of specially designed textbooks for children in Grades VII–XII. These materials reflect the new emphases and objectives in reading instruction at the secondary level that have just been cited. As the production and use of these books increase, they are destined to influence significantly the nature of reading instruction. It has appeared desirable to the present writers to attempt an appraisal of the books now available, which are listed in Table I. This task was soon found to predicate first the construction of a suitable evaluation technique. The present article describes the procedure followed in developing and formulating criteria for evaluating the books listed in Table I (and similar volumes). An accompanying article[2] in this issue of the *Journal* discusses techniques that utilize these criteria in the evaluation of reading textbooks; some basic considerations in appraising the place of such material in reading instruction at the secondary-school level also will be proposed.

PRELIMINARY CONSIDERATIONS

To be useful, a technique for evaluating reading textbooks should help in answering the following question: How adequately does the program suggested in the textbook provide for the development of appropriate reading skills and interests, and other desirable outcomes, in junior and senior high school pupils? This question raises one that is more fundamental: What are "appropriate reading skills and interests, and other desirable outcomes," for adolescent children? The latter query (about curricular objectives) must be answered first.

One valid approach to the formulation of educational objectives is through the study of the interests, abilities, and needs of the individuals for whom a program is to be planned. Accordingly, appropriate aims for a reading course may be determined after investigating these characteristics of the boys and girls of representative pupil groups in the secondary school. This method, however, involves an ambitious undertaking far beyond the scope and resources of this investigation; therefore, it was rejected.

Another possible procedure is to examine recent analyses of the needs and interests of secondary-school pupils, and to utilize the findings as a basis for building a list of aims

* On leave from Chicago Teachers College.

[1] Cf. Guy L. Bond and Eva Bond, *Developmental Reading in the High School*. New York: Macmillan Co., 1941.

[2] David Kopel, "Procedures for Evaluating Textbooks in Reading," *Journal of Experimental Education*, XII (September, 1943).

TABLE I

READING TEXTBOOKS FOR GRADES VII–XII

Title	Author	Grade Level[3]	Publisher	Date of Publication
Experiments in Reading, I, II, and III	McCall, Cook, Norvell	9–12	Harcourt-Brace	1934
Hidden Treasures in Literature, I, II, and III	McCall, Cook, Norvell	9–12	Harcourt-Brace	1934
Reading for Understanding	Bessey, Coffin	9–12	Appleton-Century	1936
Following Printed Trails	Hovious	6– 7	Heath	1936
Flying the Printways	Hovious	7– 8	Heath	1938
Reading for Skill	Broening, Law, Wilkinson, Ziegler	7–12	Noble and Noble	1936
Improving Your Reading	Wilkinson, Brown	4– 8	Noble and Noble	1938
Read and Comprehend	Knight, Traxler	9–12	Little, Brown	1937
Develop Your Reading	Knight, Traxler	7–10	Little, Brown	1941
Roads to Reading	Mack, McCall, Almack	8–10	Harcourt, Brace	1937
Self-Improvement in Reading	Pitkin, Newton, Langham	9–12	McGraw-Hill	1937
Let's Read, I	Roberts, Rand, Lundgren	7– 9	Holt	1941
Let's Read II	Roberts, Rand, Murphy, Appy	8–10	Holt	1939
Let's Read, III	Roberts, Rand	9–11	Holt	1937
Let's Read, IV	Roberts, Rand, Tardy	10–12	Holt	1940
Reading in High Gear	Cage	9–12	Harpers	1938
Quest	Herzberg, Paine, Works	7–12	Houghton Mifflin	1940
Ventures	Herzberg, Paine, Works	7–12	Houghton Mifflin	1940
Rewards	Herzberg, Paine, Works	7–12	Houghton Mifflin	1940
You and Your Reading	Stovall	7–12	Ginn	1940
Experiences in Reading and Thinking	Center, Persons	7–12	Macmillan	1940
Practices in Reading and Thinking	Center, Persons	7–12	Macmillan	1940
Problems in Reading and Thinking	Center, Persons	7–13	Macmillan	1940

[3] The grade level shown is the one for which each book was prepared, according to its author or publisher. It will be noted that a wider grade range of usefulness is claimed for some books than for others. Those with narrow ranges, of one or two grades, are designed for use with poor readers at the specified levels; they may, however, be used as developmental texts with children in lower grades. Where a wide range of grades is shown for a volume, the intent is to cover both remedial and developmental needs—remedial in the upper and developmental in the lower levels of the specified grades.

Recent additions to the list above, published after this study was completed, are: *Driving the Reading Road* and *Progress on Reading Roads*, for grades 7 and 8 respectively (Chicago: Lyons and Carnahan).

for reading instruction. Although some excellent research is being done in the field of children's interests, this alternative, also, was rejected; in this case because the results suffer from various limitations:

> The reported investigations of children's tastes and interests up to the present can not be accepted at their face value. At best, they have disclosed only broad average trends in the preferences of somewhat narrowly selected groups. Moreover, they have been concerned mainly with the compiling of choices under different conditions, seldom under conditions of adequate exposure to wide varieties of reading materials . . . Moreover, many of the studies have been so loosely controlled that the results probably do not indicate the free choices of the groups studied. Adequate techniques for determining such choices have not been discovered.[4]

A third method is to reduce the goals of various functioning programs in secondary-school reading to a composite list of commonly accepted objectives. These programs are presumably based, to some degree at least, on the needs and interests of adolescent boys and girls. This last procedure, comprehending elements of the others, was adopted as most feasible.

[4] Jean Betzner and R. L. Lyman, "The Development of Reading Interests and Tastes," *The Teaching of Reading*, p. 189. Thirty-Sixth Yearbook of the National Society for the Study of Education, Part I. Bloomington, Illinois: Public School Publishing Company. 1937.

The Objectives of Reading Programs in the Secondary School

A survey was undertaken of secondary-school reading-course objectives; these were derived from the aims suggested prominently in the textbooks themselves and in recent articles and books by such writers in the field of secondary-school reading instruction as Cross, Deals and Seamans, Eberhart, McCallister, Monroe, Paton and Pierce, Pooley, Powell, Roberts, Smith, Strang, Witty and Kopel, and Wrightstone. The consensus of opinion among these writers is that reading instruction should result in pupil improvement or growth in:

1. Comprehension
2. Vocabulary
3. Work-study skills
4. Library skills
5. Skill in reading for various purposes
6. Skill in reading different kinds of materials
7. Speed of reading
8. Efficiency of eye-movements
9. Oral reading ability
10. Tendency to read extensively
11. Familiarity with many areas of reading
12. Interest in books
13. Depth and number of worthwhile interests
14. Broadness of experience
15. Socially desirable attitudes
16. Adjustment and behavior
17. Oral and written expression
18. Ability to appreciate good writing and to distinguish between good and poor books
19. Tendency to read critically

The foregoing items, of course, are not distinct and separate entities; they merely express different facets of a dynamic matrix of good reading habits, skills, attitudes, and interests that comprise true competence in reading. The objectives suggested by this inventory were then organized and arranged as a preliminary checklist for evaluating textbooks in reading.

The next step was to restate these objectives in a form suitable for use as criteria. Thus, the objective of "improvement in comprehension" (No. 1) becomes, in criterion form, "provision for the development of comprehension." Many of the objectives were converted into criteria as easily as in the instance cited. In another instance, however, the objectives listed as "3" and "4" were considered to be very closely related and not deserving of the heavy weight they would receive by treatment as separate criteria; both objectives were therefore included in the wording of one criterion.

Considerable liberty was taken with the reformulation of some of the other objectives, in order to achieve what the writers believe to be a consistent and comprehensive set of criteria. For example, the achievement of greater "efficiency of eye-movements" (No. 8) was not believed to be an intrinsically valid objective;[5] however, some notice of this aim is appropriate and was incorporated in one criterion concerned with "provision for the correction of faulty reading habits." In another case the general and rather vague objective of "improvement in adjustment and behavior" (No. 16) was given more explicit and significant current meaning by conversion into the criterion of "emphasis on reading for democratic action." Some of the criteria, it should be noted, were derived from more than one objective. Several criteria were added that had no direct counterparts in the objectives; they express obviously necessary considerations in judging the technical construction of textbooks in reading.

The foregoing procedure yielded a checklist of twenty-five criteria; each is discussed individually below. The numbers in parentheses after each criterion are the numbers of the objectives to which the criterion is most closely related. The annotations are primarily syntheses of views expressed by various authors from whose work the criteria were derived. The writers have introduced, however, additional emphases of their own and of other contributors to the literature that seemed necessary to a description of a well-rounded and vital program of reading instruction.[6]

The Criteria

1. *Provision for development of comprehension in reading* (No. 1).—"Comprehension" refers here to skill and facility in

[5] Considerable evidence has been adduced to show that eye-movement patterns reflect reading status; they usually improve without the benefit of any specific remedial attention as other fundamental reading skills improve. *Cf.* Paul Witty and David Kopel, *Reading and the Educative Process*, pp. 138–145. Boston: Ginn and Co., 1939.

[6] The writers wish to acknowledge their indebtedness and appreciation to Professor Holland Roberts of Stanford University, who read the manuscript and whose valuable suggestions have been incorporated in several parts of the paper.

responding appropriately to printed symbols. Growth in this ability—a function in part of mental maturation and experience—is determined to an important degree by guidance or education. A reading textbook should nurture the development of this ability by frequent emphasis upon reading for meaning, and by illustrating the dependence of meaning upon context.[7]

Reading needs to be portrayed to children as a discriminative, selective process involving interaction between the reader and the writer through the medium of print. The continuity of reading and living should be shown by pointing out how the reader's background of experience influences his interpretations of print, and that the fruition of reading is to be found in modified attitudes and in action (directed, in a democratic society, toward the achievement of socially desirable human values).[8] (*Cf.* criterion No. 24.) Highly desirable and useful is the interpretation of reading as simply an indispensable, specialized form of the interpretative or discriminative process that characterizes intelligent behavior.

By conveying the foregoing and other insights about the reading process, a textbook may aid the child in understanding how he reads, and thereby help him inestimably in understanding what he reads. The various desired outcomes of reading instruction set forth in many of the following paragraphs may be considered subsidiary or closely related to this over-all objective of fostering competence in reading with understanding.

2. *Provision for development of critical reading* (No. 19).—A reading textbook should seek to dispel the notion that facts or ideas are necessarily accurate or correct simply because they appear on the printed page. It should promote a critical attitude and a healthy skepticism toward advertising claims and the editorializing indulged in by partisan newspapers. It should develop this habit of evaluation in connection with all reading materials, historical as well as contemporary. In short, it should foster the tendency to question, criticize, and evaluate the implicit as well as expressed assumptions in reading materials; and it should encourage the constant practice of checking statements by facts.[9] The good reader "not only senses what is between the lines but searches back of the words to find the purpose and watch the effects."[10]

3. *Provision for vocabulary growth* (No. 2).—This criterion refers to opportunities within the reading selections: for the acquisition of new words, for engaging in word study and etymological exercises, and for instruction in the use of the dictionary and thesaurus.

4. *Provision for development of skill in reading for various purposes* (No. 5).—A good general overview of the topic of purposes in reading is given by W. S. Gray, who mentions fourteen purposes that "have been emphasized repeatedly in recent reports and investigations."[11] The question implied in this criterion is: To what extent does the textbook emphasize and implement the idea that there are many purposes for reading, and that the reading techniques employed should be consonant with one's purposes?

5. *Provision for development of skill in reading different kinds of materials* (No. 6).—This criterion calls attention to the degree of diversity found in the selections with respect to length, literary form, and range of subjects. Here one must judge not only whether the textbook contains different kinds of materials, but also whether it provides suggestions and exercises for developing skill in reading these varied materials.

6. *Provision for development of speed in reading* (No. 7).—Suggested by this criterion are such questions as: To what extent does the textbook emphasize the importance of speed in reading? How well does it support that emphasis with speed tests, and similar devices? Does it stress that speed of reading is not a constant, but is or should be modified by one's reading purposes and by the nature of the materials being read? This point of view finds delightful expression in a statement by Clifton Fadiman:

> I do not believe dogmatically either in fast or slow reading. I believe tripe should be read practically with the speed of light and, let us say, Toynbee's *A Study of*

[7] John B. Kerfoot, *How to Read*. Boston: Houghton Mifflin Co., 1916; S. I. Hayakawa, *Language in Action*. New York: Harcourt Brace & Co., 1941.

[8] Holland Roberts, "English for United Action," *Elementary English Review*, XIX (May, 1942), 177.

[9] *Cf.* Edgar Dale, *How to Read a Newspaper*. Chicago: Scott, Foresman and Co., 1941.

[10] Holland Roberts, "Reading for Personal and Social Action," *Elementary English Review*, XIX (April, 1942), 152.

[11] W. S. Gray, Editor, *The Teaching of Reading*, pp. 33–35. Thirty-Sixth Yearbook of the N.S.S.E., Part I. Bloomington, Illinois: Public School Publishing Co., 1937.

History with tortoise deliberation. And most books are nearer to tripe than they are to Toynbee. But the trouble with practically all of us is that we suffer from chronic reverence. We make the unwarranted assumption that because a man is in print he has something to say, and, acting on this assumption, we read his every word with scrupulous care. This may be good manners, but it's a confounded waste of time.[12]

7. *Provision for improvement of oral reading* (No. 9).—This criterion suggests that a textbook should designate or include appropriate selections to be read aloud; that it should emphasize the desirability of making use of every classroom opportunity and other natural audience situations to improve oral reading ability; and that it should foster development of habits of accurate pronunciation, clear enunciation, and proper phrasing.

8. *Provision for development of work-study and library skills* (Nos. 3 and 4).—To study effectively, high-school pupils must possess the ability to read printed discourse critically and in a manner appropriate to their purpose and to the nature of the material (see criteria 1, 2, 4, 5, and 6). Needed also by these children are the somewhat technical information and skill involved in reading efficiently such non-discoursive materials as maps, tables, charts, graphs, tables of contents, indexes, menus, etc. Desirable likewise is facility in the use of card catalogues, magazine indexes, encyclopedias, and other distinctively library and reference materials. In addition, it is important to cultivate good study habits such as budgeting time and formulating work objectives. Improvement in the foregoing special study skills becomes perforce a logical part of the educational program presented in reading textbooks.

9. *Simplicity of vocabulary and language structure.*—The usefulness of a reading textbook is enhanced, if it can be read profitably by both "good" and "poor" readers of a given age or grade. This feature is present when the selections contain challenging, relatively mature ideas, and are written in simple language and readable style. (Cf. criteria 10 and 11.)

10. *Interestingness of materials.*—"Interestingness," a general quality difficult to define or measure, encompasses such factors as diversity and contemporaneity of topics; exciting, adventurous, and significant themes; freshness of treatment; and relevance to the problems and concerns of children in the junior and senior high school grades.

11. *Gradation in difficulty of reading materials.*—The importance of this criterion derives from the truism that one learns to read by reading, and from its corollary that growth in reading ability parallels the reading of increasingly difficult, challenging material. However, the new, more difficult material must not be so much more taxing than the content to which the reader is accustomed that he becomes frustrated and discouraged.

Implicit in this criterion, therefore, is the requirement that a series of books should increase in difficulty from volume to volume, and that individual volumes should advance in difficulty as one progresses from beginning to end. Unfortunately, all of the objective methods for grading the difficulty of reading material are too unreliable and cumbersome to be very useful.[13] One must depend upon subjective estimates, the validity of which depends primarily upon a wide familiarity with books and children's reactions to them.

12. *Organization according to areas of interest.*—Continuity in the reading selections of a text contributes to their readability and increases the likelihood that they will be read, if the pupil is interested in the general area represented. This criterion refers, therefore, to the grouping of selections according to the interest areas in which they belong; books are penalized when topically related selections are not arranged in some logical order or are found only at random intervals.

13. *Balance between "work-type" and "recreatory" materials* (Nos. 5, 6, 11).—The word "balance," as used here, is not meant to denote any definite mathematical relationship between the somewhat different types of reading mentioned, but serves to emphasize the desirability of substantial representation in each category. The writers recognize that the commonly expressed dichotomy between "work-type" and "recreatory" material is nonetheless arbitrary and even fictional when based solely upon the structural characteristics of the material. Whether the reading

[12] Clifton Fadiman, *Reading I've Liked*, pp. xlii–xliii. New York: Simon and Schuster, 1941.

[13] Catherine J. Elliott, "A Critical Analysis of the Objective Method of Measuring Reading Difficulty." Unpublished Doctor's Thesis, University of Pittsburgh, 1940.

of a given book or selection is "work" or "play" depends as a rule much more upon one's attitude, purpose, and skill than upon any "intrinsic" qualities of the book. The criterion calls attention then to the use of (1) materials that involve or consist of exercises, assignments, and other traditional school tasks, and (2) materials predominantly narrative or descriptive in character, and easy to read.

14. *Balance between classical and contemporary materials, and among the different literary forms* (Nos. 5, 6, 10, 11).—In this criterion, as in the one directly above, "balance" refers only to a fairly good representation of each of the types of materials indicated. It is desirable that a text introduce children to reading that is varied with respect to time, people, and problems, as well as to many styles of treatment and literary forms.[14]

15. *Amplitude of materials* (Nos. 10, 11).—While bulk in itself is obviously no virtue in a book, a small volume can not contain the many types of selections and exercises desired in a reading textbook. Amplitude of interesting, informative, worthwhile material is therefore important.

16. *Attractiveness of format.*—"Format" may be interpreted to include such factors as general appearance, size, weight, illustrations, covers, margins, type, paper, and ease of handling. These physical characteristics of a book determine in some measure whether a volume will be selected voluntarily and read with a pleasant "set" by a child. Since poor readers so often dislike books and reading, books designed for their use must look as attractive and easy to read as possible, yet withal avoid a juvenile or childish appearance.

17. *Adequacy of pupil and teacher aids.*—The textual characteristics included in this classification are: word lists, glossaries, questions, charts, drills, diagnostic tests, indices, keys, topical reading lists, manuals, and professional bibliographies.[15] These aids should be appraised for their usefulness and adequacy.

18. *Provision for correction of faulty reading habits* (No. 8).—Faulty reading habits are indicative of a retarded or immature status in reading. In other words, they are associated with or are symptomatic of poor reading. However, the perseveration of these habits acts in a vicious circle to aggravate a youngster's reading difficulties and to make them progressively worse. Fortunately, most poor reading habits, such as are reflected in faulty eye-movement patterns, improve as the child reads highly motivated and well chosen materials. Nevertheless, many undesirable habits may properly and effectively be corrected simply by describing them, by pointing out how they interfere with the achievement of meaning, and by providing the simple admonitions and exercises that help to overcome them. Included in this category of faulty habits are: excessive vocalization, lip movements, pointing, and head movements; substitutions, insertions, omissions, repetitions, and guesses in word perception; and mechanical word-calling and improper phrasing.[16]

19. *Stimulation of desire to be a good reader* (No. 12).—The introductions to practically all textbooks—and, in some volumes, additional sections—extol the values of reading. These materials attempt to stimulate an appreciation of the significant role played by reading in modern society, and of the correlative need for the individual to know how to utilize wisely this instrument of communication and medium of expression. The value of reading to the individual as an agency for enriching his personal life and for enhancing his usefulness to others should be stressed. On a lower level, added incentive for becoming a good reader is found in some texts in the provision of charts, tests, and other devices for ascertaining and recording growth in reading abilities.

20. *Stimulation of extensive reading* (Nos. 6, 10, 11).—This criterion calls attention to the desirability of encouraging pupils to read widely in many books and periodicals, and to read in relation to their interests and problems. Suggestions for supplementary reading and annotated book lists are helpful devices for implementing this objective.

21. *Stimulation of oral and written expression* (No. 17).—The arts of verbal communication—reading, writing, speaking, and listening—are closely interrelated and complementary. Common to all is much of the psy-

[14] *Cf.* Lou LaBrant, *An Evaluation of the Free-Reading in Grades 10, 11, and 12.* Contributions to Education, Vol. 1, No. 2. Columbus, Ohio: Ohio State University Press, 1936.

[15] Ivan R. Waterman and Irving R. Melbo, "A Plan of Procedure for the Evaluation of Textbooks in Reading," *Elementary School Journal*, XXXV (May, 1935), 670–671.

[16] For a more detailed statement of faulty reading habits and of methods for their elimination, see Paul Witty and David Kopel, *op. cit.*, Chs. 7 and 8.

chological process called thinking. Genuine improvement in one of these related language areas influences beneficially the status of the others, except where physical impediments interfere. Encouragement of meaningful, well-organized, purposeful oral and written expression is therefore desirable in a reading textbook as a means of fostering growth in intelligent reading.

22. *Stimulation of growth in reading tastes* (No. 18).—The kind of development desired in terms of this criterion can not be obtained in any substantial degree through a direct approach. Indeed, the common practice of telling children what is "good" and "poor" literature is best calculated to foster artificial standards and meretricious tastes. It is far more educative to encourage children to seek books that are "good" for *them*. If genuine growth is to be engendered, a child must have a series of successful reading experiences, in which he selects materials in relation to his interests and in terms of their contribution to his understandings and purposes.

23. *Emphasis on the development of desirable social attitudes* (No. 15).—Like the English course in which, as a rule, they are designed to be used, reading textbooks attempt to inculcate social attitudes and behavior traits that are commonly deemed "desirable." Among those frequently mentioned in the literature are: tolerance, sympathy, loyalty, respect for authority, obedience, thrift, persistence, patriotism, and unselfishness.

The attempt to indoctrinate attitudes and traits such as the foregoing is acceptable in a textbook, provided there is recognition that the terms used are abstractions to be interpreted not as absolutes, but as variables in relation to the situations or contexts in which they occur. For example, when respect for authority is preached without qualification and, if it is practiced blindly, without consideration of the sources of the power vested in authority and of its objectives, the consequence may be undignified, demoralizing subservience and acceptance of evil. In some situations disrespect for authority can be virtuous. Of course, experience and common sense usually correct the over-simplifications and unsound generalizations conveyed by a literal reading of much moralistic material.

The effectiveness of such material depends largely upon the form it takes. Since preaching is resented by most children, didactic and hortatory treatments should be avoided in favor of vicarious experiences that stimulate the reader to identify himself with laudable characters and discriminative, responsible conduct.

24. *Emphasis on reading for democratic action* (No. 16).—Even when part of the school day is set aside for specialized instruction in reading, teachers must not lose sight of their ongoing responsibility, transcending the demands of any "subject," to develop competent citizens for a democratic society. Our assumption is that primarily boys and girls must learn in school to live together harmoniously and to work together effectively in achieving socially desirable goals. Another premise is that children will lead the kinds of lives they learned to live as children. If their school curriculum is to equip them with the social attitudes and intellectual abilities necessary for a high level of cooperative social life, children (and teachers) at every level must be free to plan and to carry on challenging activities that represent direct participation in the work of making this a better world.

The fundamental objective of reading instruction, and indeed of education generally, has been stated with great clarity and conviction by Holland Roberts in a forthright and provocative editorial:

> Every reading act should fit into the pattern of our effort to improve the welfare of the people of the United States. The test of reading is not understanding of printed symbols, but action. The aim of the teaching of reading is good food, clothing, shelter, health, and work and recreation and the cultural enrichment of our people. If the reading program results in understanding about nutrition but does not improve the nourishment of our underfed millions, then it has failed the only important test. The well-being of the people is the final measure.[17]

25. *Encouragement of participation in various non-reading activities and experiences* (Nos. 13, 14).—The encouragement of children's participation in worthy non-reading activities serves to equip them with the meanings and purposes necessary to engage suc-

[17] Holland Roberts, "Reading for Personal and Social Action," *Elementary English Review*, XIX (April, 1942), 152.

cessfully in many types of reading. Such participation also provides vital motivation for reading that is integrated with living.

Guidance in this area is clearly a legitimate function of a reading program. Suggestions for and descriptions of socially constructive activities in which children can engage have a very important place in a reading textbook; they are especially pertinent if they stimulate related reading. Such suggestions provide the implementation for the emphasis set forth in the preceding section.[18] The value of a reading textbook may be measured roughly by its adequacy in fostering an interdependent program of useful living and reading.

Summary

Reading textbooks for Grades VII to XII represent a rather important and unique publishing venture of the last decade. Their production reflects the commendable new interest of teachers and school officials to improve the reading ability of adolescent boys and girls. The extensive and increasing use of these materials led the present writers to seek a procedure for evaluating their merits. A wide sample of the literature on reading at the secondary level was examined to obtain expert opinions as to the objectives of reading instruction. Some nineteen different aims were found to be rather commonly expressed. Statements of these aims were converted into a form suitable for use in evaluating the instructional program in reading textbooks. Other items were added to provide standards for gauging the typographical and structural adequacy of the textual materials. The resulting formulations numbered twenty-five. When read with their annotations, they are offered as a set of criteria for judging the merits of reading textbooks. The criteria and their annotations may be read also as a statement of the emphases and provisions that the writers believe should characterize the whole reading program in the upper elementary and secondary grades.

Another article will discuss two evaluational methods which utilize these criteria, as well as suggest certain other considerations that should govern the selection and employment of reading textbooks.

[18] For many valuable suggestions, see L. Rosenblatt, *Literature as Exploration*. New York: D. Appleton–Century Co., 1938.

PROCEDURES FOR EVALUATING TEXTBOOKS IN READING

DAVID KOPEL*
Army of the United States

Elsewhere is described the method employed in obtaining a set of criteria for the evaluation of reading textbooks.[1] Twenty-five criteria were formulated to represent desirable qualities in reading textbooks (and in reading programs). These criteria are as follows:

1. Provision for development of comprehension in reading.
2. Provision for development of critical reading.
3. Provision for vocabulary growth.
4. Provision for development of skill in reading for various purposes.
5. Provision for development of skill in reading different kinds of materials.
6. Provision for development of speed in reading.
7. Provision for improvement of oral reading.
8. Provision for development of work-study and library skills.
9. Simplicity of vocabulary and language structure.
10. Interestingness of selections.
11. Gradation in difficulty of reading materials.
12. Organization according to areas of interest.
13. Balance between "work-type" and "recreatory" materials.
14. Balance between classical and contemporary materials and among the different literary forms.
15. Amplitude of materials.
16. Attractiveness of format.
17. Adequacy of pupil and teacher aids.
18. Provision for correction of faulty reading habits.
19. Stimulation of desire to be a good reader.
20. Stimulation of extensive reading.
21. Stimulation of oral and written expression.
22. Stimulation of growth in reading tastes.
23. Emphasis on the development of desirable social attitudes.
24. Emphasis on reading for democratic action.
25. Encouragement of participation in various non-reading activities and experiences.

Examination of the foregoing list will indicate to most readers, perhaps, the general nature of the emphasis and bias of the criteria. Each criterion, however, contains terms such as "a good reader", terms susceptible to many diverse definitions. The meaning of some criteria can be demonstrated only through illustrations. For these reasons the writer urges any one seriously interested in the criteria to read them *with their annotations* in the accompanying article in the current issue of the *Journal*. This is imperative, if the criteria are to be applied with understanding in any evaluative procedure.

SUBJECTIVE EVALUATION

It is suggested in the companion article that a reading textbook is "good" to the degree that it possesses the qualities enumerated in the foregoing check list (and defined in the accompanying annotations of the preceding article). The criteria call attention to characteristics of a text that are significant in judging its merits. Given an understanding of the various items in the check-list, and with the list in mind or before her, a teacher can readily make an informed assay of any reading textbook, as well as a comparison of several volumes. It is thus possible to determine with some specificity whether a text is relatively good, bad, or indifferent as a whole, and in its various parts.

QUANTITATIVE EVALUATION

The foregoing procedure lends itself quite well to the evaluation of one or a few texts. It proves inadequate, however, when one is confronted with the task of making a comparative evaluation of any considerable number of books. In this situation a quantitative

* On leave from Chicago Teachers College.

[1] David Kopel and John F. O'Connor, "Criteria for Evaluating Reading Textbooks," *Journal of Experimental Education*, XII (September, 1943). Pp. 26–33.

procedure appears to be needed. Therefore, a rating device known as a scorecard was constructed.

Since there were twenty-five criteria, it was easy to use round numbers. Thus, 1000 points were made available for scoring a text that would completely satisfy all criteria. Each criterion was accorded an equal value, or 40 points. This arrangement facilitated appraisal of each of the twenty-five textual characteristics on a five-step scale: 40 points to be awarded if the text were excellent in the required characteristic, 30 points if it were good, 20 points if it were fair, 10 points if it were poor, and no points at all if it were entirely lacking in the required characteristics.[2] (The five 10 point steps permitted further subdivision if finer gradations were desired.)

A scorecard such as has been described was used in rating all the texts listed in Table I of the accompanying article.[3] Total scores for the various books ranged from 360 to 920 points.[4] Only three books, all single volumes, received less than 600 points each. The other texts, which included seven series of from two to six volumes, all received rather high total scores, in most cases between 700 and 900 points. These results indicate that most of the books possess a high degree of merit and potential usefulness.

Considerable variation, however, was found to exist among the books with respect to their excellence in meeting the standards suggested by the individual criteria. For example, of two books that obtained total scores of about 800 points, one received 0 points on "provision for development of critical reading," while the other received 40 points. Although such extreme variations in quality in any one characteristic were uncommon in books having high total ratings, this illustration does show the danger inherent in using only the total ratings for evaluation purposes.

[2] No doubt it can be argued that some criteria are more important than others, and should be weighted commensurately. However, a practical and valid method for determining proper weights did not suggest itself. Moreover, the writer reasoned (or rationalized) that whatever inequalities exist in the twenty-five items would possibly cancel each other in the application of the total scorecard. If this hypothesis is incorrect, the total scores obtained will distort the true numerical value of certain texts, exaggerating some and depreciating others. This danger provides another reason, in addition to those pointed out later in the discussion, for using total scores with great caution.

[3] It should be noted that reading textbook series (which ranged in number from two to six volumes) were considered single units for rating purposes.

[4] Ratings were made first by John F. O'Connor, a graduate student working under the writer's direction. Ratings of the same books by the writer yielded somewhat different results. Such discrepancies are to be expected, of course. However, they lead the writer to refrain from publishing detailed scores until they can be validated by obtaining and pooling the ratings of a number of qualified judges. Any interested teacher, school official, or psychologist, not employed by a publishing house or associated with the preparation of the volumes listed in Table I, is invited to cooperate in this project. Upon receipt of an inquiry the writer will furnish such persons with scorecard and directions for rating the reading textbooks. Judges must furnish their own books and they must agree to read all of the books carefully before rating them. The ratings received during a three-month period will be averaged and the data, if substantial, will be published.

Other Considerations in Evaluating Textbooks

It is clear that the total scores derived from applying the scorecard to the reading textbooks should not be considered absolute measures of their merits. Of greater significance in determining the value of a text for a given group of children are the ratings of the book on each criterion or characteristic. Moreover, each book and each characteristic of every book should be evaluated in terms of its usefulness for a specific child, i.e., in terms of the relevance of the materials to the interests and needs of the individual. Such considerations may lead to the selection of a book that ranks low on the scorecard; nevertheless, it may be the best book for a particular youngster at a given time. Children's needs thus become the paramount considerations in selecting reading textbooks for them. Moreover, since the children in any class will display diverse interests and needs, a teacher will rarely meet these needs by using any single volume as a class text; she will do far better to regard these special volumes as manuals or reference books to be used only at appropriate times by children whose special needs they meet.

A desirable procedure is to inventory the needs of a group and then to select as many (or more) books as there are children in the group; the books will be chosen finally in terms of their correlation with children's needs. Frequently one finds small groups within a class with more or less common problems; for them one provides a dozen copies of one title, seven or eight of another, and perhaps half a dozen of a third book. For the remaining individuals one purchases single copies, thus extending the variety and range of materials available to the whole class. A typical library in a reading classroom would then include at least one volume of nearly all of the available books. The purchase of additional copies of any book

would be determined by the evaluation procedures and consideration of children's needs as described.

The Place of Texts in a Reading Program

Although this report implies and, indeed, stresses the value of texts, it should not be inferred necessarily from this discussion that reading textbooks constitute the only or even the most significant component of a sound reading program. Experts hold somewhat divergent views as to the proper role of textbooks in classroom instruction. For example, Strang suggests certain values in their use:

> For gifted well-prepared teachers with extensive library facilities practice books are unnecessary, for they can plan with each pupil individually a valuable progression of reading experiences. For overburdened teachers with little background in the theory and procedures of reading instruction, the practice books provide suggestions and materials which they can use to good advantages.[5]

Witty, on the other hand, questions "the desirability of recommending a basic text, the use of which will undoubtedly bring about some improvement in present conditions." He believes that: "It is doubtful whether such a recommendation (were it to be followed) would represent a forward and beneficial step".[6] His reasons are:

> . . . Genuine and enduring improvement in reading will occur, . . . only in so far as teachers become expert in studying and understanding children and in directing and guiding them so that worthy individual patterns of growth in reading will develop. Teachers, therefore, must become serious and competent students of children's growth; and in addition, they must acquire thorough acquaintance with children's literature, old and new. There are more productive ways of teaching reading than by use of a text; free-reading programs and experience curriculums have abundantly demonstrated this. One citation only, Lou LaBrant's *An Evaluation of a Free Reading Program*, is evidence enough. The issue here is patently clear: Are we, as teachers, interested predominantly in a body of subject matter which all children will study; or are we concerned primarily with other objectives, viz., the preservation of individuality, personality orientation and adjustment, and continuous sturdy growth. A position on this question will determine our interest in and use of basic texts.[7]

It should be noted that Witty argues against the use of such books as we have been discussing as *basic* texts. In other words, he opposes a mechanical, inflexible reading program that is oriented with primary reference to a prescribed series of exercises contained in some practice book. With this position the present writer concurs, but holding to this position does not, in the writer's opinion, preclude the *proper* use of reading textbooks. Expressed earlier was the view that: (1) not one basic text, but several texts should be employed in a typical class; (2) these books should be selected and used only in so far as they relate to definitely ascertained children's needs; (3) reading textbooks should be regarded not as "basic texts", but rather as reference books, manuals, workbooks, etc.; and (4) they should be considered to constitute one element only in a well-rounded, abundant classroom library of diversified reading materials.

No single book is a substitute for a rich library and for a teacher who understands children and books, and how to bring them together. With this limitation of texts thoroughly understood, teachers in Grades VII–XII need have no reluctance in providing a place for many reading textbooks in their developmental, corrective, and remedial reading classes. Even in free-reading programs and experience curriculums, some children will probably continue, at least in the immediate future, to display reading deficienies that may be corrected in part through judicious use of selected materials from reading textbooks. For some years to come, therefore, the reading textbook seems destined to provide a valuable supplement to other essential facilities of a sound reading program.

[5] Ruth Strang, *Problems in the Improvement of Reading in High School and College*, p. 282. Lancaster, Pennsylvania: Science Press Printing Co., 1938.

[6] Paul A. Witty, "A Review of *Growth in Reading*," *English Journal*, XXVIII (May, 1939), 407.

[7] *Loc. cit.*

READING REACTIONS FOR VARIED TYPES OF SUBJECT MATTER

An Analytical Study of the Eye Movements of Eighth Grade Pupils[1]

EARL WILLIAM SEIBERT
Guidance Director and Psychologist
Belleville Schools, New Jersey

THE PROBLEM AND ITS IMPORTANCE

The purpose of this study was two-fold: (1) to determine the nature of the eye movements of eighth-grade pupils while reading different types of subject matter and to identify relationships that may exist between eye movements and various types of subject matter during the reading process, and (2) to determine the effect that various parts of a selection have on eye movements as subjects progress through a 300 word reading selection.

Reading problems hold a prominent place in the minds of present-day educators. Various methods have been used to analyze problems of reading difficulty. One of these has been the study of photographic records of the eye movements of individuals while they are reading. It has been established that efficiency in reading is reflected in oculomotor behavior.

Early investigators in the field of oculomotor behavior were concerned primarily with what their studies of eye movements would reveal about the reading process. Their studies dealt with few subjects. However, beginning in 1917, investigators became more concerned about the individual differences to be found in the records of eye movements. New developments made it possible to secure longer records from each subject and to use more individuals. It was found that there were varied and complex patterns of eye movements among individual subjects. The individual varies his eye movements to meet the particular reading situation at hand. Eye-movement habits are flexible and appear to adjust readily to any change in the perceptual processes involved.

Recent investigators have shown the need for additional research in eye movements when the individual reads different types of subject matter and when he reads longer selections than have customarily been used in eye-movement investigations. There is also the need for intensive studies of the eye-movement habits of individuals at a particular age level. No intensive study of the eye movements of eighth-grade boys and girls has been reported.

The present investigation is a study of the eye movements of sixty eighth-grade boys and girls while reading various types of subject matter. The selections of subject matter are academic material of average eighth-grade level. These selections are 300 words in length and have been equated on the basis of vocabulary difficulty and of number and length of sentences.

EXPERIMENTAL PROCEDURE

Eighth-grade pupils in the public schools of Belleville, New Jersey, cooperated as subjects in the experiment. Belleville school pupils are typical of the boys and girls of the nation in their performance on certain objective tests. The thirty-one boys and twenty-nine girls selected as subjects are a representative sample of the eighth-grade pupils in Belleville.

The photography of eye movements was done in January and February of 1942 by the investigator, who is guidance director and psychologist for the board of education, Belleville, New Jersey.

Paragraph selections were taken from academic material of six different types: (1) mathematics, (2) biography, (3) adventure, (4) physical science, (5) history, and (6) geography. Juries of New York University professors or comparable public-school leaders selected typical eighth-grade subject matter for each of the six types. Each selection was 300 words in length. The six selections finally accepted were comparable in number and length of sentences and in vocabulary difficulty.

[1] Abstract of a study submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the School of Education of New York University, 1943.

The selections were printed on non-glare index cards in 10-point type. The length of the lines was 3.5 inches and two-point leading was used between the lines.

The Ophthalm-O-Graph, a portable binocular eye-movement camera, was used for photographing the eye movements as the subject read the material from a card attached to the instrument. A standard uniform procedure was used in giving the directions to the subjects throughout the tests. Each subject read an orientation selection of fifty words before beginning the reading of the longer selections. Two selections were administered at each setting with at least one day intervening between sittings.

A comprehension test of from fourteen to twenty-four questions was prepared for each selection. These questions were answered by the subject immediately after the reading of each selection. The subject was told before reading the selection that he would be asked to answer questions on it. The validity of each test item was established before the items were used in this experiment, in a preliminary study conducted by the investigator with the 135 pupils in the eighth grade of Derry Township public schools, Hershey, Pennsylvania. The Hershey and Belleville eighth-grade pupils were comparable. Only significant items were used in the scoring of the comprehension tests.

Results

The purpose of the first phase of this study was to determine the nature of the eye movements of eighth-grade pupils while reading different types of subject matter and to identify relationships that may exist between eye movements and various types of subject matter during the reading process.

The mean number of fixations per 100 words was found for the sixty subjects on each of the six types of subject matter. The reliability of the difference in the number of fixations required to read 100 words was computed between the various selections. The number of regressions per 100 words, the average duration of fixation (in seconds), the average span of recognition (in words), the rate of reading in words per minute, and the comprehension scores were given the same statistical treatment. Critical ratios are shown in Table I.

The range within each measure and for each selection was found. To show the variation with which the same individual read the different selections, standard deviations were computed for the performance of each individual on all selections. This was done for each eye-movement measure.

The average comprehension score of the individual for the six selections was correlated with the language part of the California Test of Mental Maturity and with the average for each of the five eye-movement measures. The various eye-movement measures were correlated with each other. The average comprehension score for all the subjects in Belleville was compared with the average comprehension score for all of the Hershey subjects.

Significant differences on all the five eye-movement measures are evident in Table I. The differences are greater for comprehension than they are for any of the five eye-movement measures. The differences in the type of selection seem to have more effect on the comprehension of the selection than they have on the eye movements used during the comprehension of the selection.

There are variations in the eye-movement measures when the same individual reads different selections. There are individuals who read all the selections with approximately the same number of fixations; with other individuals the difference in the selection is accompanied by a difference in the number of fixations. The same is true of regressions, of span of recognition, and of rate of reading. In each case there were individuals who read all the selections with variation in that particular eye-movement measure, but there were also many subjects who read all the selections with little variation in that particular eye-movement measure. This was not true of duration of fixation; there is little variation in the duration of fixation a subject uses when reading the different selections.

There is little relationship between comprehension scores and eye-movement measures. Number of fixations, number of regressions, duration of fixation, span of recognition, or rate of reading can not be predicted on the basis of the scores in comprehension. Reading comprehension and eye-movement measures are two different phases of the complex process of reading.

TABLE I

CRITICAL RATIOS BETWEEN SELECTIONS ON ALL MEASURES
Ratio of the Difference to the σ of the Difference

Identification of Selection		Fixations	D/σD favors	Regressions	D/σD favors	Duration of Fixation	D/σD favors	Span of Recognition	D/σD favors	Rate of Reading	D/σD favors	Comprehension	D/σD favors
A	Mathematics												
B	Biography	0.50	B	0.66	B	0.53	A	0.62	B	0.71	A	2.88	A
A	Mathematics												
C	Adventure	1.21	A	1.93	A	2.31	A	1.23	A	1.90	A	2.92	A
A	Mathematics												
D	Physical Science	0.75	A	1.90	A	2.53	A	0.14	A	1.80	A	0.18	A
A	Mathematics												
E	History	1.16	A	1.96	A	2.41	A	0.86	A	2.23	A	2.27	A
A	Mathematics												
F	Geography	1.11	A	0.45	A	1.35	A	0.40	A	1.27	A	2.01	A
B	Biography												
C	Adventure	1.70	B	2.43	B	1.90	B	1.76	B	1.25	B	0.07	B
B	Biography												
D	Physical Science	1.24	B	2.42	B	2.08	B	0.68	B	1.12	B	2.62	D
B	Biography												
E	History	1.65	B	2.50	B	1.97	B	1.41	B	1.56	B	0.37	E
B	Biography												
F	Geography	1.62	B	1.00	B	0.86	B	0.96	B	0.62	B	0.46	F
C	Adventure												
D	Physical Science	0.45	D	0.15	D	0.07	D	0.94	D	0.17	D	2.66	D
C	Adventure												
E	History	0.06	E	0.16	E	0.16	E	0.35	E	0.24	C	0.44	E
C	Adventure												
F	Geography	0.13	F	1.35	F	1.08	F	0.78	F	0.59	F	0.52	F
D	Physical Science												
E	History	0.39	D	0.00	neither	0.10	E	0.62	D	0.42	D	2.05	D
D	Physical Science												
F	Geography	0.33	D	1.27	F	1.15	F	0.21	D	0.44	F	1.82	D
E	History												
F	Geography	0.06	F	1.30	F	1.06	F	0.43	F	0.85	F	0.10	F

When investigations are made of eye movements, as much attention should be given to the construction of adequate tests to measure the comprehension of the material read as is granted to the securing of an accurate photographic film record.

The photography of eye movements does not need to create an unnatural reading situation for the subject. If proper rapport is established, pupils can be expected to give a valid record of their reading habits.

Means on the six measures are shown in Table II for each of the different selections.

The group means for all of the selections on each of the eye-movement measures were: fixations, 127.00 ± 3.19 per 100 words; regressions, 29.58 ± 1.21 per 100 words; duration of fixation, .2837 ± .00331 seconds; span of recognition, .837 ± .0214 words; and rate of reading, 182.93 ± 5.91 words per minute.

The purpose of the second phase of this study was to determine the nature of the eye movements of the subjects in reading various parts of a 300 word selection.

The mean number of fixations per 100 words, based on the averages of all the subjects on all the selections, was found for the first fifty words, the first hundred words, the second hundred words, the third hundred words, and the entire 300 words. The reliability of the differences in average number of fixations between each of the parts of a 300 word selection was computed. The reliability of the difference between the number of fixations for the first fifty words and the first hundred words was computed for each selection. This was also done for the first hundred words and the second hundred words, and for the second hundred words and the third hundred words.

The same statistical treatment was given to the data of the number of regressions per 100 words, the average duration of fixation (in seconds), the average span of recognition (in words), and the rate of reading in words per minute.

When the means for each of the five parts of a selection are based on all the six selections, there is a trend in the direction of greater efficiency of eye movements during the reading of the third hundred words than during the reading of any other part of a 300 word selection. The next greatest efficiency

TABLE II

MEANS ON SIX MEASURES FOR DIFFERENT SELECTIONS

		Selections					
		A	B	C	D	E	F
Fixations (per hundred words)	M	123.83	121.50	129.83	127.50	129.50	129.17
	σ_M	3.33	3.24	3.67	3.59	3.59	3.45
	σ	25.80	25.12	28.42	27.77	27.81	26.78
	σ_σ	2.35	2.28	2.59	2.53	2.53	2.44
Regressions (per hundred words)	M	27.75	26.67	31.50	31.17	31.17	28.58
	σ_M	1.11	1.20	1.59	1.42	1.34	1.47
	σ	8.63	9.29	12.30	10.98	10.39	11.36
	σ_σ	.79	.85	1.12	1.00	.95	1.03
Duration of Fixation (in seconds)	M	.2738	.2767	.2877	.2873	.2868	.2813
	σ_M	.00403	.00363	.00448	.00351	.00357	.00385
	σ	.0312	.0285	.0347	.0272	.0277	.0298
	σ_σ	.00280	.00260	.00316	.00248	.00252	.00271
Span of Recognition (in words)	M	.845	.865	.805	.840	.817	.832
	σ_M	.0214	.0239	.0243	.0281	.0243	.0246
	σ	.166	.185	.187	.218	.187	.191
	σ_σ	.0151	.0168	.0170	.0198	.0170	.0174
Rate of Reading (words per minute)	M	194.17	188.32	177.50	179.00	175.42	182.93
	σ_M	5.96	5.76	6.43	6.01	5.95	6.57
	σ	46.17	44.62	49.80	46.57	46.07	50.88
	σ_σ	4.20	4.06	4.53	4.24	4.19	4.63
Comprehension (per cent)	M	44.59	35.17	34.89	44.09	36.68	37.10
	σ_M	1.89	2.67	2.73	2.12	2.94	3.22
	σ	14.63	20.67	21.17	16.44	22.81	24.96
	σ_σ	1.33	1.88	1.93	1.50	2.08	2.27

TABLE III

CRITICAL RATIOS BETWEEN THE MEANS FOR EACH OF THE FIVE PARTS OF A SELECTION, BASED ON SCORES OF 60 SUBJECTS ON ALL THE SELECTIONS

Ratio of the Difference to the σ of the Difference

	Part of Selection	Fixations	$D/\sigma D$ favors	Regressions	$D/\sigma D$ favors	Duration of Fixation	$D/\sigma D$ favors	Span of Recognition	$D/\sigma D$ favors	Rate of Reading	$D/\sigma D$ favors
A	1st 50 words										
B	1st 100 words	.18	A	.58	A	1.07	A	1.11	A	1.15	A
A	1st 50 words										
C	2nd 100 words	.48	C	.53	C	2.00	A	.59	A	.92	A
A	1st 50 words										
D	3rd 100 words	.98	D	1.70	D	1.54	A	.36	D	.44	A
A	1st 50 words										
E	Entire 300 words	.34	E	.54	E	1.40	A	.49	A	1.04	A
B	1st 100 words										
C	2nd 100 words	.65	C	1.10	C	.92	B	.54	C	.25	C
B	1st 100 words										
D	3rd 100 words	1.13	D	2.25	D	.47	B	1.51	D	.69	D
B	1st 100 words										
E	Entire 300 words	.51	E	1.11	E	.29	B	.69	E	.10	E
C	2nd 100 words										
D	3rd 100 words	.47	D	1.15	D	.45	D	.97	D	.45	D
C	2nd 100 words										
E	Entire 300 words	.15	C	0.00	neither	.67	E	.12	E	.15	C
D	3rd 100 words										
E	Entire 300 words	.63	D	1.16	D	.20	E	.89	D	.59	D

of eye-movement measures takes place during the reading of the first fifty words. (See Table III).

The rate of efficiency of eye-movement measures established during the reading of the first fifty words is decreased during the reading of the second fifty words; the reader re-orients himself during the second hundred words and finally attains his greatest efficiency of eye-movement measures during the reading of the third hundred words. This is true when the data are based on the reading of all the selections.

When the data are considered on the basis of individual selections, there is greater efficiency of eye-movement measures during the first fifty words than during the first hundred words. There are no significant differences in the efficiency with which the first hundred and second hundred words are read. There is greater efficiency during the third hundred words than during the second hundred words when the data are considered for individual selections.

Therefore, whether the data are computed in terms of all the six selections or in terms of each individual selection, the pattern of eye movements of eighth-grade pupils during the reading of a 300 word selection is the same; the rate of efficiency of eye-movement measures established during the reading of the first fifty words is decreased during the reading of the second fifty words; the reader re-orients himself during the second hundred words and finally attains his greatest efficiency of eye-movement measures during the reading of the third hundred words (Table IV).

Conclusions

1. There are significant differences for each of the five eye-movement measures in the reading of selections of different types of subject matter.

2. The differences between selections are greater for comprehension than they are for any of the eye-movement measures.

3. There are variations in eye-movement measures for individual subjects, as well as for groups of subjects.

4. There is little relationship between comprehension scores and eye-movement measures; this is to be expected, for both are different phases of the complex process of reading.

5. Both an eye-movement film record and an adequate test of comprehension on the material read are important parts of a reading examination.

6. If proper rapport is established, pupils can be expected to give a valid eye-movement film record of their reading habits.

7. The first fifty words of a 300 word selection give a fair sample of the eye movements of eighth-grade individuals, but a more accurate picture of the pattern of eye-movement measures is provided by the film record of the third hundred words.

Implications of Findings

1. Each subject teacher in the eighth grade needs to be a teacher of reading.

2. Some eighth-grade pupils maintain a definite pattern of eye-movement habits in the reading of different types of subject matter, but most eighth-grade pupils vary their eye-movement habits when reading selections of different types.

3. Possibilities for the increasing of the span of recognition and for the training of other eye-movement habits need to be explored at various age levels.

4. The teaching of vocabulary and emphasis on techniques of getting the thought from the printed page should aid the reading comprehension of the pupils and therefore play a part in the improvement of the reading process.

Table IV

The Various Means Based on All the Selections

	1st 50	1st 100	2nd 100	3rd 100	Entire 300	
Fixations	128.50	129.33	126.33	124.17	127.00	per 100 words
Regressions	30.50	31.50	29.58	27.58	29.58	
Duration of fixation	.2770	.2823	.2870	.2847	.2837	seconds
Span of recognition	.853	.815	.833	.865	.837	words
Rate of reading	191.68	182.08	184.17	187.91	182.93	words per minute

5. If an individual pupil has desirable eye-movement habits but inadequate comprehension, emphasis on vocabulary and the techniques of getting thought from the printed page may aid in the improvement of the reading process.

6. If an individual pupil reads with good comprehension, but has undesirable eye-movement habits, it is possible that development of more desirable eye-movement habits may aid in the improvement of the reading process.

7. A film record of a fifty-word selection will give a fair picture of the eye-movement habits of an individual. A film record of a fifty-word selection in each of several types of subject matter will indicate the extent to which an individual pupil varies and adjusts his eye-movement habits to meet the demands of a different selection. When a very complete picture of the reading habits of an individual is required, a 300 word selection in each of several types of subject matter will need to be read with eye-movement photography being done on the third hundred words.

Bibliography

1. Alderman, Everett. "The Effect of Size of Type on Speed of Reading and the Determination of Various Factors That May Influence the Results." Doctor's Thesis, University of Pittsburgh, 1938.
2. Bender, Irving E., Imus, Henry A., Rothney, John W. M., Kemple, Camilla, and England, Mary R. *Motivation and Visual Factors.* Hanover, N. H.: Dartmouth College Publications, 1942.
3. Betts, Emmett A. *The Prevention and Correction of Reading Difficulties.* Evanston, Illinois: Row, Peterson and Company, 1936.
4. Blair, Glenn Myers. "Instrumentation and the Remedial Reading Program," *Educational Administration and Supervision,* XXVI (1940), 201–208.
5. Broom, M. E. "The Reliability of the Reading Graph Yielded by the Ophthalmograph," *School and Society,* LII (1940), 205–208.
6. Broom, M. E., and Others. *Effective Reading Instruction.* New York: McGraw-Hill Book Company, 1942.
7. Buswell, Guy T. *Fundamental Reading Habits: A Study of Their Development.* Supplementary Educational Monographs, No. 21. Chicago: University of Chicago, 1922.
8. Buswell, Guy T. *How Adults Read.* Supplementary Educational Monographs, No. 45. Chicago: University of Chicago, 1937.
9. Cattell, James McKeen. "The Inertia of the Eye and Brain," *Brain,* VIII (1885), 295–312.
10. Center, Stella S., and Persons, Gladys L. *Teaching High School Students to Read.* New York: D. Appleton–Century Company, 1937.
11. Dearborn, Walter F. *The Psychology of Reading.* Columbia University Contributions to Philosophy, Psychology, and Education, Vol. XIV. New York: Columbia University, 1906.
12. Delabarre, E. B. "A Method of Recording Eye-Movements," *American Journal of Psychology,* IX (1898), 572–574.
13. Dodge, Raymond. "Reaction Time of the Eye," *Psychological Review,* VI (1899), 477–483.
14. Dodge, Raymond. "Visual Perception During Eye Movement," *Psychological Review,* VII (1900), 454–465.
15. Dodge, Raymond, and Cline, Thomas S. "The Angle Velocity of Eye-Movements," *Psychological Review,* VIII (1901), 145–157.
16. Duke-Elder, W. Stewart. *Text-Book of Ophthalmology,* Vol. I. London: Henry Kimpton, 1932.
17. Eurich, Alvin C. "The Reliability and Validity of Photographic Eye-Movement Records," *Journal of Educational Psychology,* XXIV (1933), 118–122.
18. Gilbert, L. C., and Gilbert, D. W. "Reading Before the Eye-Movement Camera versus Reading Away From It," *Elementary School Journal,* XLII (1942), 443–447.
19. Gray, William S. *Reading in General Education.* Washington: American Council on Education, 1940.
20. Gray, William S. *Recent Trends in Reading.* Supplementary Educational Monographs, No. 49. Chicago: University of Chicago, 1939.

21. Huey, Edmund B. *The Psychology and Pedagogy of Reading*. New York: Macmillan Company, 1908.

22. Imus, H. A., Rothney, J. W. M., and Bear, R. M. *An Evaluation of Visual Factors in Reading*. Hanover, N. H.: Dartmouth College Publications, 1938.

23. Judd, Charles H. *Reading: Its Nature and Development*. Supplementary Educational Monographs, No. 10. Chicago: University of Chicago, 1918.

24. Judd, Charles H., and Buswell, Guy T. *Silent Reading: A Study of the Various Types*. Supplementary Educational Monographs, No. 23. Chicago: University of Chicago, 1922.

25. Miles, W. R., and Shen, Eugene. "Photographic Recording of Eye Movements in the Reading of Chinese in Vertical and Horizontal Axes," *Journal of Experimental Psychology*, VIII (1925), 344–362.

26. Paterson, Donald G., and Tinker, Miles A. "Influence of Line Width on Eye Movements," *Journal of Experimental Psychology*, XXVII (1940), 572–577.

27. Robinson, F. P. "Speed versus Comprehension in Reading — A Discussion," *Journal of Educational Psychology*, XXXI (1940), 554–558.

28. Schmidt, William A. *An Experimental Study in the Psychology of Reading*. Supplementary Educational Monographs, No. 2. Chicago: University of Chicago, 1917.

29. Sisson, E. Donald. "The Causes of Slow Reading: An Analysis," *Journal of Educational Psychology*, XXX (1939), 206–214.

30. Slocum, Roger Leon. "Reading Status of University of Wisconsin Freshmen." Unpublished Ph. D. Thesis, University of Wisconsin, 1940.

31. Stone, Lewis Gordon. "Reading Reactions for Varied Types of Subject Matter: An Analytical Study of the Eye-Movements of College Freshmen," *Journal of Experimental Education*, X (September, 1941), 64–77.

32. Taylor, Earl A. *Controlled Reading*. Chicago: University of Chicago Press, 1937.

33. Tinker, Miles A. *A Photographic Study of Eye Movements in Reading Formulae*, pp. 68–182. Genetic Psychology Monographs, Vol. III, No. 2, 1928.

34. Tinker, Miles A. "Eye Movements in Reading," *Journal of Educational Research*, XXX (1936), 241–277.

35. Tinker, Miles A. "Trends in Diagnostic and Remedial Reading as Shown by Recent Publications in this Field," *Journal of Educational Research*, XXXII (1938), 293–303.

36. Tinker, Miles A. "The Role of Eye Movements in Diagnostic and Remedial Reading," *School and Society*, XXXIX (1934), 147–148.

37. Tinker, Miles A. "Speed versus Comprehension in Reading as Affected by Level of Difficulty," *Journal of Educational Psychology*, XXX (1939), 81–94.

38. Traxler, Arthur E. *Ten Years of Research in Reading*. Educational Records Bulletin, No. 32. New York: Educational Records Bureau, 1941.

39. Vernon, M. D. *The Experimental Study of Reading*. Cambridge: Cambridge University Press, 1931.

40. Witzeman, B. Evangeline. "An Experimental Study Using the Ophthalm-O-Graph and Metron-O-Scope in the Diagnosis and Treatment of Reading Defects," *Journal of Psychology*, XI (1941), 307–334.

THE PLOT-COMPLETION TEST

For Use in High Schools by Guidance Counselors, School Psychiatrists, and Teachers of Literature*

SARAH I. RODDY

PURPOSE OF THE TEST

This test was constructed for the purpose of analyzing the pupil's attitude toward fictitious situations and, indirectly by implication, toward his own life problems. It furnishes a crude measurement of certain specific abilities: first, to grasp all the facts in a given situation; second, to evaluate suggested solutions in the light of the given facts and the pupil's knowledge of life as gained from experience and observation; and, third, to select a true-to-life solution. In general, it yields interesting implications as to whether the pupil's approach to life problems is scientific and well integrated. Such an analysis of the pupils' reactions will be useful to school psychiatrists, psychologists, guidance counselors, home room teachers, and all other persons interested in the personality development of school children. It will also give considerable assistance to the teacher of literature, for it will not only help him to understand the members of his class, but will also enable him to proceed with greater insight in guiding their reading and interpretation of literature.

Some attitudes that may warp the pupil's thinking about literature and about life are morbidity, wishful thinking in its various forms, and superstition. The test is designed to reveal such tendencies on the part of persons answering the questions.

DESCRIPTION OF THE TEST

The *Plot-Completion Test* presents ten "plot germs", or situations involving conflict. For each plot five endings are listed, which the pupil is asked to number in the order of their probability as outcomes of the given situation. The most true-to-life ending is to be labeled 1, the one next in probability 2, and so on. If the pupil completes the test before the close of the class period, he is asked to go back through the plots and star for each one the ending that he thinks would make the best story, irrespective of whether it is true to life or not. Pupils who still have time after doing this are asked to check thus (√) for each plot the ending that describes the most advisable course of action. (These additional responses, included partly to engage the attention of quick workers who are likely to grow restless before their classmates finish the test, are also interesting as indications of whether the pupil recognizes the fact that the ending he prefers may not be the one most likely to occur.)

Some specific attitudes that may prevent the pupil's proper evaluation of the solutions are: (1) the desire for a happy ending, even if it necessitates a lucky accident, a sudden and unwarranted change in character, or a violation of the laws of nature; (2) the desire to reward the likable person and punish the less amiable one, regardless of who is in the right in the given situation; (3) a morbid preference for an unhappy outcome, even when a cheerful one is more probable; (4) a tendency to reject the given problem and substitute an easier one; and (5) a superstitious willingness to credit mystical mental powers and other unscientific beliefs. These five attitudes are significant indications of the pupil's personality and of his taste in reading. The maker of the test included them because she had been noticing for several years the frequency with which they appeared in original short stories and plays written by members of her English classes and in their comments about their reading. She had observed also that certain types of classroom behavior seemed to parallel these attitudes toward fictitious situations. The test is designed to reveal these attitudes where they exist to a marked degree. Since a person may think emotionally about some problems encountered in fiction and objectively about others, according to the degree of their relevancy to his own case, each plot in the test is based upon an event that might conceivably happen to the pupil; and the situations are planned to include relationships involving parent and

* Developed as part of a doctorate dissertation completed at New York University, 1943.

child, teacher and pupil, friends of the same sex, and friends of the opposite sex, in order to cover a sufficiently wide range to reveal the pupil's emotional bias, if he has a very marked one.

It is granted that some pupils whose reactions to their own life situations are affected by such attitudes as the test is meant to reveal may not evidence these attitudes in completing the plots, because the test is presented as an intellectual task and the problems are therefore artificial. It is easier to solve another person's problems than one's own, and still easier to solve hypothetical than real ones, for in these cases one's point of view is more objective. Despite this fact, it frequently happens that the pupil identifies himself with some of the characters in the plots, particularly if his thinking is habitually colored by his emotions. Moreover, the very fact that the test is presented to the pupil as a measurement of literary ability, rather than as a personality questionnaire, somewhat lessens the wariness of the sophisticated taker of personality tests. In testing approximately 400 high school seniors the maker of this test has found relatively few whose answers gave no evidence of any of the attitudes that the test is meant to disclose.

How the Plot-Completion Test Was Developed

Before the final version of the test was reached, two successive preliminary forms were tried out on various groups of high school seniors in Rockland County, New York. A large number of high school teachers, graduate students in the School of Education of New York University, psychological experts, and miscellaneous adults also took the tests. Everyone who answered the questions was invited to comment on them. The comments guided the maker of the test in revising it. While these comments revealed certain weaknesses in the early forms of the test, they also confirmed the examiner in her impression that a form could be developed that would achieve her purpose. When the shortcomings revealed by this preliminary investigation had been eliminated, the examiner deleted all questions on which no pupil had made any mistakes and all those on which all pupils had made mistakes, and rearranged the remaining ones in order of their difficulty. *The Plot-Completion Test* is the final form resulting from the investigation.

The Scoring Key

The scoring key was based upon the judgment of ten psychological specialists, who ranked the endings for each plot in order of their probability. They submitted their answers individually, without comparing notes. To serve as a check against the answers of the ten experts, ten members of the English Education Department in the School of Education of New York University, 25 student nurses at the University of Portland, and 18 miscellaneous educated adults chosen at random from the acquaintances of the examiner took the test. Their answers and those of 253 Rockland County high school seniors were tabulated and compared with those of the ten specialists. The specialists' scoring key was corroborated by the consensus of all groups.

Credits were assigned for answers in such a way as to indicate as accurately as possible every gradation in discrimination between realistic and unrealistic thinking.

Reliability

In order to measure the reliability of the test, the examiner correlated the odd and even scores of the 319 Rockland County high school pupils who had up to that time taken the test. By the split-half method it was learned that the coefficient of reliability of half the test is .7168. Using the Spearman–Brown formula, the examiner found that the coefficient of reliability of the entire test is .8350. The index of reliability was found to be .9138. In other words, merely temporary circumstances have comparatively little effect upon the score of a pupil taking this test, and the score he obtains is very likely to be the true measure of his ability in the areas being tested.

Validity

Since the index of reliability of a test is a measure of its theoretical validity, the *Plot-Completion Test* is clearly a valid criterion of that which it directly tests—the pupil's ability to predict the outcomes of situations, avoiding certain pitfalls offered by appeals to wishful thinking and by other tricks frequently employed in pseudo-realistic fiction. Evidence that the test fulfills this purpose is

also present in the comments of the ten psychological experts who served as judges for the test.

A further check on the usefulness of the test in revealing the pupil's success or failure in detecting the specified falsities in plot was furnished by an analysis of a set of original short stories written by a group of high school pupils who also took the *Plot-Completion Test*. In order to compare the implications of the short stories with those of the test answers, the examiner selected from the stories all those containing marked instances of the falsities in plot that the test is meant to measure. Each of these stories (18 in number) was analyzed in comparison with the same pupil's test paper. The purpose of the comparison was to ascertain whether the pupils who had written unrealistic stories had also received low scores in the test, and whether the same types of errors had been made by each pupil in both of his papers. A marked similarity appeared in the two papers of 12 of the 18 pupils. In five other cases there was a general similarity, and in only one case was there no similarity at all.

For the reasons explained above, it may be assumed that the *Plot-Completion Test* has validity in measuring ability to discriminate between realistic and pseudo-realistic plots. Whether it also gauges validly a pupil's adjustment to life is quite another matter. Therefore, the examiner correlated the scores of 91 students on Victor Noll's *What Do You Think? A Test in Scientific Thinking*[1] with their scores on the *Plot-Completion Test*. The coefficient of correlation was found to be between .14 and .78 on the one percent level of certainty. For predicting scores of individuals from one test to the other, this degree of correlation would be too low, but it is a dependable indication that some correlation exists between the two tests. Therefore, since the validity of Noll's test has been established,[2] it is reasonable to conclude that a pupil's answers to the *Plot-Completion Test* may yield some valuable *implications* as to his approach to life problems. More than that this examiner does not claim for her test. She believes that even this amount of help in diagnosing individual personalities will be useful.

Uses of the Plot-Completion Test

Because of its reliability and validity, this test may be used for diagnostic purposes and to measure the improvement made by classes in their ability to detect some of the tricks of pseudo-realism. Such information would be of service to both the teacher of literature and the guidance counselor.

The examiner believes that her test is even more useful as a teaching device than as a measure of improvement. Pupils who have taken the test are always eager to discuss it, and such discussion may be made very fruitful. It offers a specific starting point for simple lessons in mental hygiene. It may also motivate instruction on the differences between books that present a true picture of life and those which only purport to do so.

[1] Victor H. Noll, *What Do You Think? A Test of Scientific Thinking*. New York: Bureau of Publications, Teachers College, Columbia University, 1935. The test is accompanied by a handbook for teachers, called *The Habit of Scientific Thinking*.

[2] As evidence of the validity of his test, Noll has shown that the scores are little dependent upon native intelligence, that the separate items of the test consistently discriminate between persons making high total scores and those making low total scores on the test, and that this discrimination agrees with the trend of the total scores. See pages 23–25 of the handbook.

INFLUENCE OF STUDENT DREADS UPON ATTITUDES TOWARD SCHOOL SUBJECTS

WINONA M. PERRY
University of Nebraska

What are the attitudes or feelings that accompany the early achievement of pupils who have entered a class either with anticipation, or with dread lest they be unable to cope with its subject matter? Will such like, indifference, or dislike: (1) persist throughout their study of this course, and (2) develop into a strong desire to have the opportunity to study more widely or more intensively in this field?

These questions, together with the findings of previous experimental attacks upon this problem, were incorporated into subject questionnaires, one for each subject studied on the high school or the college level. During 1940–41 these questionnaires were presented to students in the University of Nebraska enrolled in introductory courses in educational psychology and in high school methods. In general, these students, as will be observed in Tables Ia, Ib, and Ic, were university sophomores or juniors (84%) whose experiences on the high school level occurred mainly in small or large city schools (52%) or in town or consolidated schools (47%).[1] Ninety-eight percent of these students stated that the average evaluation of their scholastic achievement in high school was a mark of A or B, but that in college their present status would be described more nearly by a mark of B or C.

The frequencies of the three categories of students who liked, or were indifferent toward, or disliked a given subject revealed that the proportions of students in each category were similar, whether the students had studied in one of the larger or smaller high schools, were then upper or lower classmen, possessed a psychological rating of B or C, or had achieved a scholastic level of B or C. Thus, no one of these factors could be regarded as a major cause in developing student attitudes of liking or indifference toward or dislike of a given subject. Consequently the investigator turned to a critical examination of another contributing factor—the influence of dreading or not dreading one's experiences in a subject area before starting to study that subject.

Sources of data were the responses of students to sixteen subject-questionnaires. Of these, nine pertained to subjects that these students had studied while in high school, and seven to college subjects which they had studied or were studying at that time. Summaries of these data appear in Figures I and II. The legend, which describes the frequency totals of answers to the several questions as tabulated for English composition, will guide the reading of each frequency tabulation for any subject studied in high school (English, American history, world history, algebra, geometry, foreign languages, general science, biology, and chemistry) or in college (English composition, English literature, American history, foreign languages, biology, chemistry, and geography). Although the total number of students was 225, the maximum number of responses to any one item on any subject-questionnaire was 221.

[1] For one of the groups that included 124 students, ratings on the A.C.E. Psychological Examination (1940 Edition) were available. Eighty of these students (64%) earned ratings that indicated student ability to reach at least a B or C achievement level in college. (See Table Id.)

SUBJECTS DREADED, OR NOT DREADED

The order in which subjects on both high school and college levels were dreaded is:

High School Subject[2]		Frequency of Dread	College Subject[2]	
English	7%	Least	6%	English literature
General science	11%			
Biology	14%	Less often	----------	
World history	15%			
American history	18%			

[2] See Table II regarding percentages of students reporting in each subject.

High School Subject[2]		Frequency of Dread		College Subject[2]
Algebra	22%	Average	21%	English composition
Foreign Languages	23%		23%	American history
			24%	Foreign languages
			25%	Geography
Chemistry	31%	More often	35%	Biology
Geometry	35%			
		Most often	55%	Chemistry

[2] See Table II regarding percentages of students reporting in each subject.

Therefore, among these sixteen subjects investigated, college chemistry holds a unique position, in that it is the one subject which was dreaded by more than fifty percent of its students. Geometry and chemistry in high school join with biology and chemistry in college in being those subjects that students regard with the greatest degree of apprehension. While, at the other extreme, English in high school and English literature in college emerge as those subjects that students approach, not with hesitation, but with a more positive attitude of anticipating their experiences in these fields.[3]

The statement of reasons for dreading or not dreading to study these subjects, it will be noted from Figures I and II, have much smaller frequencies, owing to the fact that all students were urged to state their reasons only when known, and according to a given classification, with or without elaboration by illustration or explanation. However, in general, the most potent reason for dreading or not dreading a subject is the nature of its subject matter. For all high school subjects except mathematics, the teacher is the greater, or approximately equal, cause for dreading or not dreading a subject than is one's estimate of his ability, the pattern being S(Subject)>T(Teacher)>Y(Yourself); but, for college subjects, this finding is reversed, since one's self is stated as the cause more, or as frequently, than is the teacher (S>Y≧T). There are two exceptions to this general trend and each appears on the high school level: (1) among the "not-dreads" students, both the nature of the subject matter and the teacher seem to be about equally effective causes of attitudes toward geometry, foreign languages, biology, and chemistry (S≧T>Y); (2) but those students who had dreaded to study algebra and geometry regard their own inability as the cause more frequently than the teacher (S>Y>T).

[3] An attempt was made to discover whether the fact that a subject is required or elective in a curriculum, or whether the number of semesters which had been devoted to its study, would be indicative of the frequency of dread of that subject. The data, however, show that as yet neither fact is predictive of the position which a subject will occupy on this "scale of frequency of dread".

Attitude Toward Subjects

What are some of the concomitant attitudes accompanying achievement in these school subjects? Will certain subjects attract and others repel our students? The findings of Table IVa, derived from data presented in Figures I and II and in Table III, may be summarized thus:

Students		Exceptions
All	L>I≧D	Foreign languages, geography (college)
"Not dreads"	L>I≧D	
"Dreads"	D>(I or L)	Biology, chemistry, Eng. comp. (college)

In general, when students had completed a semester of study in these subjects, their liking for these subjects far outweighed their attitudes of indifference or of dislike, although these latter attitudes occurred with approximately equal frequency in two college courses, foreign languages and geography. However, although this conclusion remains true among the "not-dreads", those students who had dreaded these subjects had, on the whole, developed dislike of them more frequently than either liking of or indifference toward them. But, among the "dreads", on the college level, dislike, indifference, and like appeared in biology and chemistry, and in part in English composition, with about equal frequency; while on the high school level, frequency of student dislike of geometry outnumbered that of like and of indifference by a ratio of two to one, and their dislike of foreign languages occurred twice as often as did that of indifference, and four times that of like.

FIGURE I*

HIGH SCHOOL SUBJECTS

ENGLISH

	Not Dr.	Dr.
Required	145	13
Elective	59	1

Semesters of Study	Not Dr.	Dr.
8	130	9
7	21	0
6	33	2
5	8	1
2, 3, 4,	9	4

STATUS:

Initial	Not Dr.	Dr.	Final L		Final I		Final D	
L	162	3	147	3	13	0	2	0
I	30	5	8	1	21	3	1	1
D	2	7	0	1	1	3	1	3
Total	194	15	155	5	35	6	4	4

DESIRE TO ELECT Another Course

	Not Dr.	Dr.
Yes	131	3
Ind.	31	7
No	32	0
Intensity	90	8

REASONS:

	Those who had (or had not) dreaded to start studying it: Not Dr.	Dr.
S(Subject)	22	7
T(Teacher)	14	6
Y(Yourself)	7	5

Like, Indifference, or Dislike Persisting or Changing throughout Course

	From L Not Dr.	Dr.	From I Not Dr.	Dr.	From D Not Dr.	Dr.
S	65	2	9	0	0	3
T	41	0	6	2	2	3
Y	11	1	1	2	0	0

HISTORY: AMERICAN

	Not Dr.	Dr.
Required	138	35
Elective	19	0

Semesters of Study	Not Dr.	Dr.
4	7	1
3	1	1
2	122	30
1	27	5

STATUS:

Initial	Not Dr.	Dr.	Final L		Final I		Final D	
L	131	4	123	2	4	2	4	0
I	26	8	5	2	21	5	0	1
D	3	24	1	6	0	4	2	14
Total	160	36	129	10	25	11	6	15

DESIRE TO ELECT Another Course

	Not Dr.	Dr.
Yes	86	9
Ind.	34	5
No	31	21
Intensity	63	17

REASONS:

	Those who had (or had not) dreaded to start studying it: Not Dr.	Dr.
S(Subject)	24	27
T(Teacher)	14	11
Y(Yourself)	7	9

Like, Indifference, or Dislike Persisting or Changing throughout Course

	From L Not Dr.	Dr.	From I Not Dr.	Dr.	From D Not Dr.	Dr.
S	51	1	7	4	1	8
T	38	2	6	1	1	10
Y	7	0	1	1	0	1

HISTORY: WORLD

	Not Dr.	Dr.
Required	81	16
Elective	42	15

Semesters of Study	Not Dr.	Dr.
5, 6	1	0
4	5	0
3	5	0
2	96	18
1	21	4

STATUS:

Initial	Not Dr.	Dr.	Final L		Final I		Final D	
L	98	0	87	0	8	0	3	0
I	27	8	6	2	19	4	2	2
D	3	14	0	4	0	4	3	6
Total	128	22	93	6	27	8	8	8

DESIRE TO ELECT Another Course

	Not Dr.	Dr.
Yes	69	3
Ind.	28	3
No	32	16
Intensity	51	8

REASONS:

	Those who had (or had not) dreaded to start studying it: Not Dr.	Dr.
S(Subject)	17	17
T(Teacher)	11	6
Y(Yourself)	4	5

Like, Indifference, or Dislike Persisting or Changing throughout Course

	From L Not Dr.	Dr.	From I Not Dr.	Dr.	From D Not Dr.	Dr.
S	45	0	13	5	0	6
T	27	0	7	1	2	7
Y	7	0	1	0	1	3

* See end of Figure II for explanatory note on how to read this Figure.

FIGURE I—Continued

HIGH SCHOOL SUBJECTS

ALGEBRA

	Not Dr.	Dr.
Required	129	37
Elective	32	6

Semesters of Study	Not Dr.	Dr.
5, 6	2	1
4	34	4
3	35	6
2	86	29
1	5	5

STATUS:

Initial	Not Dr.	Dr.	Final L		Final I		Final D	
L	127	9	108	7	12	0	7	2
I	26	10	5	3	20	7	1	0
D	2	26	1	4	0	7	1	15
Total	155	45	114	14	32	14	9	17

DESIRE TO ELECT Another Course

	Not Dr.	Dr.
Yes	68	6
Ind.	34	5
No	53	31
Intensity	63	15

REASONS:

	Those who had (or had not) dreaded to start studying it: Not Dr.	Dr.
S(Subject)	22	34
T(Teacher)	18	9
Y(Yourself)	4	18

Like, Indifference, or Dislike Persisting or Changing throughout Course

	From L: Not Dr.	Dr.	From I: Not Dr.	Dr.	From D: Not Dr.	Dr.
S	57	2	9	3	4	9
T	28	3	7	0	2	9
Y	15	2	4	4	0	4

GEOMETRY

	Not Dr.	Dr.
Required	85	53
Elective	37	6

Semesters of Study	Not Dr.	Dr.
4	2	1
3	5	3
2	110	54
1	8	8

STATUS:

Initial	Not Dr.	Dr.	Final L		Final I		Final D	
L	87	7	77	7	4	0	6	0
I	31	11	6	0	22	9	3	2
D	5	47	0	10	1	7	4	30
Total	123	65	83	17	27	16	13	32

DESIRE TO ELECT Another Course

	Not Dr.	Dr.
Yes	47	6
Ind.	21	5
No	56	52
Intensity	45	28

REASONS:

	Those who had (or had not) dreaded to start studying it: Not Dr.	Dr.
S(Subject)	15	48
T(Teacher)	16	15
Y(Yourself)	3	20

Like, Indifference, or Dislike Persisting or Changing throughout Course

	From L: Not Dr.	Dr.	From I: Not Dr.	Dr.	From D: Not Dr.	Dr.
S	36	7	13	3	3	22
T	24	6	8	1	2	11
Y	8	1	2	2	0	3

LANGUAGES: FOREIGN

	Not Dr.	Dr.
Required	67	25
Elective	78	14

Semesters of Study	Not Dr.	Dr.
7, 8	8	2
5, 6	19	4
3, 4	97	30
2	23	10
1	6	0

STATUS:

Initial	Not Dr.	Dr.	Final L		Final I		Final D	
L	125	9	105	3	11	3	9	3
I	19	11	2	3	16	4	1	4
D	4	25	0	1	0	4	4	20
Total	148	45	107	7	27	11	14	27

DESIRE TO ELECT Another Course

	Not Dr.	Dr.
Yes	73	6
Ind.	22	1
No	51	37
Intensity	70	25

REASONS:

	Those who had (or had not) dreaded to start studying it: Not Dr.	Dr.
S(Subject)	16	30
T(Teacher)	15	12
Y(Yourself)	9	10

Like, Indifference, or Dislike Persisting or Changing throughout Course

	From L: Not Dr.	Dr.	From I: Not Dr.	Dr.	From D: Not Dr.	Dr.
S	50	5	10	4	2	15
T	27	4	3	4	2	6
Y	11	3	0	3	0	1

FIGURE I—Continued

HIGH SCHOOL SUBJECTS

GENERAL SCIENCE

	Not Dr.	Dr.
Required	70	10
Elective	27	2

Semesters of Study	Not Dr.	Dr.
4	2	0
3	1	0
2	72	10
1	28	3

STATUS:

Initial \ Final	Not Dr.	Dr.	L		I		D	
L	69	0	61	0	6	0	2	0
I	29	6	7	1	20	3	2	2
D	2	7	0	1	0	2	2	4
Total	100	13	68	2	26	5	6	6

DESIRE TO ELECT Another Course

	Not Dr.	Dr.
Yes	51	2
Ind.	23	3
No	27	8
Intensity	45	3

REASONS:

	Those who had (or had not) dreaded to start studying it: Not Dr.	Dr.
S(Subject)	16	9
T(Teacher)	12	6
Y(Yourself)	5	0

Like, Indifference, or Dislike Persisting or Changing throughout Course

	From L: Not Dr.	Dr.	From I: Not Dr.	Dr.	From D: Not Dr.	Dr.
S	30	1	13	4	1	3
T	11	0	12	3	0	2
Y	3	0	3	0	0	1

BIOLOGY

	Not Dr.	Dr.
Required	42	9
Elective	33	2

Semesters of Study	Not Dr.	Dr.
4	1	0
3	1	0
2	71	9
1	13	4

STATUS:

Initial \ Final	Not Dr.	Dr.	L		I		D	
L	68	6	62	4	6	1	0	1
I	14	6	1	1	11	4	2	1
D	0	2	0	0	0	1	0	1
Total	82	14	63	5	17	6	2	3

DESIRE TO ELECT Another Course

	Not Dr.	Dr.
Yes	46	3
Ind.	16	1
No	24	7
Intensity	31	5

REASONS:

	Those who had (or had not) dreaded to start studying it: Not Dr.	Dr.
S(Subject)	17	9
T(Teacher)	16	3
Y(Yourself)	8	2

Like, Indifference, or Dislike Persisting or Changing throughout Course

	From L: Not Dr.	Dr.	From I: Not Dr.	Dr.	From D: Not Dr.	Dr.
S	26	2	6	2	0	0
T	24	2	6	4	1	2
Y	2	1	2	1	0	1

CHEMISTRY

	Not Dr.	Dr.
Required	11	6
Elective	28	12

Semesters of Study	Not Dr.	Dr.
4	0	1
3	1	1
2	40	16
1	2	1

STATUS:

Initial \ Final	Not Dr.	Dr.	L		I		D	
L	28	6	23	5	2	0	3	1
I	13	3	2	1	9	2	2	0
D	2	10	1	4	0	0	1	6
Total	43	19	26	10	11	2	6	7

DESIRE TO ELECT Another Course

	Not Dr.	Dr.
Yes	19	5
Ind.	7	2
No	18	12
Intensity	19	10

REASONS:

	Those who had (or had not) dreaded to start studying it: Not Dr.	Dr.
S(Subject)	8	16
T(Teacher)	7	3
Y(Yourself)	5	6

Like, Indifference, or Dislike Persisting or Changing throughout Course

	From L: Not Dr.	Dr.	From I: Not Dr.	Dr.	From D: Not Dr.	Dr.
S	15	3	6	2	1	5
T	8	2	4	0	2	4
Y	5	0	0	0	1	2

FIGURE II

COLLEGE SUBJECTS

ENGLISH COMP.	Not Dr.	Dr.
Major	26	6
Minor	53	15
Group Req.	94	25
Elective	2	0
Semesters of Study	Not Dr.	Dr.
5, 6	7	1
4	12	5
3	52	12
1, 2	100	26

STATUS:

Initial \ Final	Not Dr.	Dr.	L		I		D	
L	123	4	99	3	15	1	9	0
I	44	21	9	6	30	15	5	0
D	8	21	3	1	1	3	4	17
Total	175	46	111	10	46	19	18	17

DESIRE TO ELECT Another Course

	Not Dr.	Dr.
Yes	88	6
Ind.	36	5
No	34	30
Intensity	57	21

REASONS:	Those who had (or had not) dreaded to start studying it: Not Dr.	Dr.
S(Subject)	13	33
T(Teacher)	7	11
Y(Yourself)	6	27

Like, Indifference, or Dislike Persisting or Changing throughout Course

	From L: Not Dr.	Dr.	From I: Not Dr.	Dr.	From D: Not Dr.	Dr.
S	46	3	11	10	3	10
T	37	1	17	6	3	6
Y	11	1	5	4	3	3

ENGLISH LITERATURE	Not Dr.	Dr.
Major	31	1
Minor	41	2
Group Req.	42	5
Elective	10	1
Semesters of Study	Not Dr.	Dr.
5, 6	5	0
3, 4	23	0
2	49	4
1	41	5

STATUS:

Initial \ Final	Not Dr.	Dr.	L		I		D	
L	103	0	85	0	6	0	12	0
I	14	2	3	1	9	0	2	1
D	0	5	0	1	0	0	0	4
Total	117	7	88	2	15	0	14	5

DESIRE TO ELECT Another Course

	Not Dr.	Dr.
Yes	89	1
Ind.	10	0
No	20	5
Intensity	65	6

REASONS:	Those who had (or had not) dreaded to start studying it: Not Dr.	Dr.
S(Subject)	12	7
T(Teacher)	4	2
Y(Yourself)	4	3

Like, Indifference, or Dislike Persisting or Changing throughout Course

	From L: Not Dr.	Dr.	From I: Not Dr.	Dr.	From D: Not Dr.	Dr.
S	42	0	5	2	0	4
T	27	3	5	0	0	3
Y	8	0	1	1	0	0

HISTORY: AMERICAN	Not Dr.	Dr.
Major	11	2
Minor	23	3
Group Req.	27	14
Elective	2	0
Semesters of Study	Not Dr.	Dr.
4, 5	3	1
3	4	0
2	34	9
1	22	9

STATUS:

Initial \ Final	Not Dr.	Dr.	L		I		D	
L	47	6	40	5	3	1	4	0
I	18	8	4	1	11	4	3	3
D	0	5	0	0	0	5	0	0
Total	65	19	44	6	14	10	7	3

DESIRE TO ELECT Another Course

	Not Dr.	Dr.
Yes	35	6
Ind.	9	6
No	18	7
Intensity	34	8

REASONS:	Those who had (or had not) dreaded to start studying it: Not Dr.	Dr.
S(Subject)	9	11
T(Teacher)	4	3
Y(Yourself)	2	8

Like, Indifference, or Dislike Persisting or Changing throughout Course

	From L: Not Dr.	Dr.	From I: Not Dr.	Dr.	From D: Not Dr.	Dr.
S	26	5	7	3	0	2
T	7	1	4	5	1	0
Y	3	1	1	1	0	1

FIGURE II—Continued

COLLEGE SUBJECTS

LANGUAGES:

FOREIGN	Not Dr.	Dr.
Major	7	0
Minor	19	5
Group Req.	31	14
Elective	25	11

Semesters of Study	Not Dr.	Dr.
7, 8	3	0
5, 6	9	2
3, 4	38	9
1, 2	32	17

STATUS:

Initial \ Final	Not Dr.	Dr.	L		I		D	
L	69	9	57	6	4	2	8	1
I	10	1	0	0	6	1	4	0
D	2	16	0	3	1	1	1	12
Total	81	26	57	9	11	4	13	13

DESIRE TO ELECT Another Course

	Not Dr.	Dr.
Yes	49	4
Ind.	8	2
No	25	22
Intensity	40	16

REASONS:

	Those who had (or had not) dreaded to start studying it: Not Dr.	Dr.
S(Subject)	14	21
T(Teacher)	6	8
Y(Yourself)	4	15

Like, Indifference, or Dislike Persisting or Changing throughout Course

	From L: Not Dr.	Dr.	From I: Not Dr.	Dr.	From D: Not Dr.	Dr.
S	30	3	4	1	2	8
T	11	2	3	1	2	7
Y	9	0	1	0	0	3

BIOLOGY:

	Not Dr.	Dr.
Major	4	3
Minor	9	1
Group Req.	35	23
Elective	7	2

Semesters of Study	Not Dr.	Dr.
4	2	1
3	1	0
2	24	17
1	27	10

STATUS:

Initial \ Final	Not Dr.	Dr.	L		I		D	
L	44	3	33	3	4	0	7	0
I	7	10	1	1	4	8	2	1
D	2	16	0	7	0	2	2	7
Total	53	29	34	11	8	10	11	8

DESIRE TO ELECT Another Course

	Not Dr.	Dr.
Yes	20	3
Ind.	7	6
No	23	20
Intensity	26	9

REASONS:

	Those who had (or had not) dreaded to start studying it: Not Dr.	Dr.
S(Subject)	12	26
T(Teacher)	6	5
Y(Yourself)	4	8

Like, Indifference, or Dislike Persisting or Changing throughout Course

	From L: Not Dr.	Dr.	From I: Not Dr.	Dr.	From D: Not Dr.	Dr.
S	19	3	3	6	2	8
T	9	4	4	3	2	4
Y	2	1	0	0	0	3

CHEMISTRY:

	Not Dr.	Dr.
Major	4	6
Minor	5	2
Group Req.	16	25
Elective	7	1

Semesters of Study	Not Dr.	Dr.
5, 6	1	0
3, 4	6	4
2	17	18
1	8	13

STATUS:

Initial \ Final	Not Dr.	Dr.	L		I		D	
L	18	8	12	4	3	4	3	0
I	8	7	1	4	5	3	2	0
D	3	21	2	3	1	4	0	14
Total	29	36	15	11	9	11	5	14

DESIRE TO ELECT Another Course

	Not Dr.	Dr.
Yes	16	7
Ind.	3	6
No	12	21
Intensity	13	14

REASONS:

	Those who had (or had not) dreaded to start studying it: Not Dr.	Dr.
S(Subject)	6	30
T(Teacher)	3	8
Y(Yourself)	3	19

Like, Indifference, or Dislike Persisting or Changing throughout Course

	From L: Not Dr.	Dr.	From I: Not Dr.	Dr.	From D: Not Dr.	Dr.
S	6	6	3	2	1	8
T	5	2	5	1	2	2
Y	2	3	3	2	2	4

FIGURE II—Continued

COLLEGE SUBJECTS

GEOGRAPHY:

	Not Dr.	Dr.
Major	2	2
Minor	0	0
Group Req.	39	12
Elective	6	1

Semesters of Study	Not Dr.	Dr.
7, 8	0	1
3	1	0
2	28	8
1	19	5

STATUS:

Initial	Not Dr.	Dr.	Final L		Final I		Final D	
L	35	1	23	1	1	0	11	0
I	9	1	2	0	6	0	1	1
D	1	13	0	4	0	0	1	9
Total	45	15	25	5	7	0	13	10

DESIRE TO ELECT Another Course

	Not Dr.	Dr.
Yes	15	1
Ind.	9	2
No	21	12
Intensity	23	6

REASONS:

	Those who had (or had not) dreaded to start studying it: Not Dr.	Dr.
S(Subject)	8	15
T(Teacher)	2	4
Y(Yourself)	0	5

	Like, Indifference, or Dislike Persisting or Changing throughout Course From L: Not Dr.	Dr.	From I: Not Dr.	Dr.	From D: Not Dr.	Dr.
S	18	2	4	0	2	7
T	12	1	2	0	1	3
Y	1	0	0	0	2	1

Legend: English Composition (college)

	Not Dr.	Dr.
Major	26	6
Minor	53	15
Group	94	25
Elective	2	0

Of those students **who had not dreaded** to study this subject, 26 reported this subject as their major, 53 as their minor, and 94 as a course included in a "group requirement" toward a degree earned in this university. Similarly, among those **who had dreaded** to study this subject, the numbers of student reports are 6, 15, and 25 respectively.

Semesters of Study	Not Dr.	Dr.
5, 6	7	1
4	12	5
3	52	12
1, 2	100	26

Among the "not dread" students, seven studied this subject five (or six) semesters, and 12, 52, and 100 studied it four, three, and one (or two) semesters respectively. Among the students who had dreaded the study of this subject, the frequencies are one, five, twelve, and twenty-six respectively.

	Not Dr.	Dr.	FINAL L		I		D	
L	123	4	99	3	15	1	9	0
I	44	21	9	6	30	15	5	0
D	8	21	3	1	1	3	4	17
Total	175	46	111	10	46	19	18	17

There were 123 students who had not dreaded to study this subject and who liked their beginning work in this study. Although 99 continued to like to study it, 24 changed their attitude (15 becoming indifferent, and nine changing to actual dislike). Among the four "dreads", three continued to like it, one became indifferent, and no one changed toward dislike of this subject. This table likewise presents the numbers of students ("not dreads" and "dreads"): (1) who initially were indifferent toward this subject and (2) those who disliked it initially; each is followed by the frequencies of those who persisted in or who changed their attitudes toward this subject.

DESIRE TO ELECT Another Course	Not Dr.	Dr.
Yes	88	6
Ind.	36	5
No	34	30
Intensity	57	21

Many students (94) indicated their preferences. These preferences follow: 88 "not dreads" and six "dreads" did desire to elect another course in this field; but 41 (36 "not dreads" and five "dreads" were indifferent, and 64 (34 "not dreads" and 30 "dreads") preferred not to study further in this field That such desires were intense reactions, was indicated by 78 students (57 "not dreads" and 21 "dreads").

REASONS:

(1) Forty-six votes (13 "not dreads" and 33 "dreads") pointed to the nature of the subject matter of English composition as the chief cause of students who had, or who had not dreaded to start their study of this subject; likewise 18, and 33 votes, indicated the cause as being their teacher, or themselves, respectively.

(2) Forty-nine students (46 "not dreads" and three "dreads") who liked this subject initially and who continued to like it (or who changed to indifference, or to dislike) attributed their attitude chiefly to the nature of its subjectmatter. But 38 students similarly listed the teacher and 12 regarded their own abilities or inabilities as the main cause of the persistence, or change, of their attitudes throughout their study of this subject. Similarly, we may note the causes (S, T, Y) for persistence or change of attitude (L, I, D) on the part of those students whose initial status had been that of indifference, or of dislike.

Persistence of Attitudes

Do these attitudes persist or change as the student progresses in a subject during a semester of study? Examination of Figures I and II yields facts of which the following are illustrative on the high school and college levels:

	Attitude:	Persisted	Changed From L	Changed To L
English (H. S.)	Like	150	15	10
	Indifference	24	11	17
	Dislike	4	5	4
Chemistry (Col.)	Like	16	10	10
	Indifference	8	7	12
	Dislike	14	10	5

On the whole, student liking of a subject predominated initially and persisted throughout the course. However, there was an accompanying current of activity as students changed from one attitude to another.

Comparatively, which subjects revealed greater persistence or change in attitudes developed initially and during the semester? Comparisons of frequencies of all student likes with their non-likes, obtained through combination of the frequencies of indifference and of dislike, are expressed in ratios (Table IVb). These ratios indicate the approximate number of times that the frequency of *like*(L) exceeded the frequency of *non-like*(I plus D). These findings emerge:

Independence of Attitudes

Let us examine another approach to this question of possible association between these two factors. In general, are the student attitudes of not dreading or dreading a given course independent of their liking or non-liking that course? Or, is there a positive relationship between these attitudes? A technique for studying this question is described by Fisher.[5] Collecting all frequencies in the two classifications of "not-dreads and dreads" and of "like and non-like" from the end status data for any one subject, e.g., geometry from Figure I, we have the following 2 x 2, or fourfold table:

	Not Dr.	Dr.	Totals
Like	83	17	100
Non-Like	40	48	88
Totals	123	65	188

[5] R. A. Fisher, *Statistical Methods for Research Workers*, pp. 88–95. Edinburgh and London: Oliver and Boyd, (Seventh Edition), 1938.

	HIGH SCHOOL	COLLEGE
1. "Scaled order" in which subjects appeared:		
Subjects **liked most frequently**, initially and finally:	English Biology American history World history Algebra General science Geometry Chemistry	English literature Foreign languages American history
Subjects **equally often liked and non-liked**, initially and finally: Ratios: 2 : 1		English composition Biology Chemistry
Ratios: 1 : 1	----------	
Subjects **more frequently non-liked than liked**, initially and finally: Ratios: 2 : 1		
2. Comparison of "scaled orders":		
Subjects **liked less frequently finally than initially**:[4]	English Biology Foreign languages	English literature Foreign languages Geography
3. Order of subject fields in college compared with that in high school:	INITIALLY	FINALLY
Subjects **liked more frequently** in college:	English literature Foreign languages	Foreign languages
Subjects **liked equally often** in college and high school: Ratios: 2 : 1 and 2 : 1	American history	American history
Subjects **liked less frequently** in college than in high school:	English composition Biology	English composition Biology

[4] These lists change somewhat when we examine all decreases of six or more, in the frequencies of student liking from the beginning to the end of the semester on the part of each group of students, "not dreads" and "dreads". The changes are:

	HIGH SCHOOL	COLLEGE
"Not Dreads" Added:	Algebra	Biology, English composition
"Dreads": Substituted:	History (American and World) Geometry	Biology, English composition

Applying the formula for Chi square[6] to this fourfold table, we have:

$$X^2 = \frac{(83\cdot48 - 17\cdot40)^2 \cdot 188}{100\cdot88 \cdot 123\cdot65} = 29.1701$$

Since the "number of degrees of freedom" is one, this value of X^2 (29.17) corresponds to the probability[7] that P>.01. Thus, dreading or not dreading to study geometry and liking or non-liking it by the end of the semester are clearly not independent attitudes. In other words, it does make a difference whether a pupil does or does not dread geometry before he starts to study it.

Values of X^2 with the corresponding values of the probability (P) for each subject on the high school and college levels are stated in Tables Va and Vb. The fourfold tables pertaining to high school chemistry, college chemistry, and geography do not indicate that we can reject the hypothesis of independence between the factors of dreading(or not) and later liking(or not) these subjects. In all other subjects, including eight on the high school level and five on the college level, we have found a significant difference between these student attitudes, that of dreading(or not) and their later liking(or not) the subject studied. The value of X^2, however, does "not measure the degree of association" between these attitudes;[8] to measure the latter, it would be necessary to inquire beyond the data available in this study.[9]

Causes of Persistence of Attitudes

Various reasons[10] for development and persistence or change of student attitudes of like, or indifference, or of dislike of these school subjects indicate that, as indicated in Figures I and II:

1. Unlike the reasons for dreading or not dreading a subject preceding their study of it, students in high school and college agree not only that the most potent cause for liking a subject and for such liking to persist, or in a few cases to change to indifference or to dislike, is the nature of the subject itself; but their statements likewise point, though to a much lesser degree, to the teacher as the more potent contributing cause than is their estimate of their own ability to master the subject matter, the pattern being S>T>Y.

2. When student reaction is indifference, whether such persists or changes to like or dislike of a subject, the contributing factors for high school pupils agree mainly with those reported under their liking of a subject, although the number of cases and the differences between the reported causes are much smaller. College students, however, who had not dreaded their approaching study of a subject, while agreeing in American history, report that their indifferent attitudes toward chemistry and English composition are due mainly to the teacher and only in decreased emphasis to the nature of the subject and estimates of their own abilities, (T>S≧Y). The latter reflects their attitude toward all other subjects except one, namely, that those who had dreaded chemistry regard the teacher as the least, and the nature of the subject and their own ability as the equally greater contributing causes.

3. High school pupils regard the teacher as the most important reason for their dislike of English and of American and world history, but for all other subjects they name the nature of the subject matter; the number of cases, however, is small. For college students, the nature of the subject matter and the teacher are stated most frequently as the contributing causal factors (S≧T>Y), with the exception of chemistry, in which they rate their own ability or inability as a more potent factor than the teacher. (S>Y>T).

In summary, the reaction of high school and college students toward the factors contributing to the development of their attitudes is in general found to be, in decreasing order: the nature of the subject matter, the teacher, and their estimates of their own ability to master this subject matter.

Motivating Power of School Subjects

To what extent did study of these school subjects arouse the pupils' desires to elect another course(s) in the same subject field? From examination of Figure III, we may draw the following conclusions:

All students.—If left to their own desires, students while in high school would wish to select a further course in English. This is

[6] *Ibid.*, p. 90.

[7] *Ibid.*, Table of X^2, pp. 118–19.

[8] *Ibid.*, p. 94.

[9] If, e.g., we could assume that amounts of dread, and similarly of liking of a given subject are distributed according to a normal distribution, tetrachoric *r*'s could reveal the degree of association between these factors. But the data presented in Figures I and II would not warrant our acceptance of this assumption.

[10] These are subject to the limitations described under the examination of reasons for dreading a school subject.

true, likewise, for college students of English literature. In high school some students would not wish to elect courses in algebra, foreign languages, or chemistry, and their opposition toward further study of geometry becomes much more frequent and definite. The latter similarly is descriptive of their reaction toward the college subjects of chemistry, biology, and geography.

"Not-dreads".—The reaction on the part of those students who had not dreaded the study of these subjects parallels the above findings quite closely, but with these exceptions: (1) the more specific opposition of the students toward additional courses in high school chemistry and algebra, and (2) a somewhat more favorable attitude toward each subject on the college level.

FIGURE III

COMPARISON (EXPRESSED IN RATIOS)[11] BETWEEN NUMBERS OF STUDENTS WHOSE DESIRES WOULD, OR WOULD NOT, DIRECT THEM TOWARD FURTHER STUDY IN THESE SUBJECTS

All Students

	HIGH SCHOOL SUBJECTS	COLLEGE SUBJECTS
WOULD (YES)	3.	3. Eng. Lit.
	2. English	2.
Neutral (Y=N)	1. Am. Hist., World Hist., Gen. Sci., Biol. Alg., For. Languages, Chem.	1. Eng. Comp., Am. Hist., For. Languages
WOULD NOT (NO)	2.	2. Chemistry
	Geometry	Biology
	3.	3. Geography

STUDENTS NOT DREADING / STUDENTS DREADING[12] — THESE SUBJECTS

WOULD (YES)	3.	3. Eng. Lit.
	2. English	2.
	Am. History Biol.	Eng. Comp., Am. Hist., For. Langs.
Neutral (Y=N)	1. World Hist., For. Langs., Gen. Sci., Alg., Geom., Chem.	1. Chemistry
		Biology
WOULD NOT (NO)	2.	2. Geography (Am. History)
	(English, Biology)	
	3. (Am. History, Chemistry)	3.
	4.	4. (Chemistry)
	5.	5. (Eng. Lit.)
	(Gen. Science)	
	6. (Algebra)	6. (Eng. Comp., For. Langs.)
	(World History, For. Langs.)	
	7.	7.
	8.	8.
		(Biology)
	9.	9.
	(Geometry)	
	10.	10.
	14.	14. (Geography)

[11] These ratios, found in Table V, have been expressed to the nearest integer.

Legend: H. S. subjects for all students:

English is the only subject in which the greater number of pupils have indicated their wish to study another course in the same subject. Their desire is expressed by the ratio of nearly 2:1.

American history, world history, general science, and biology have approximately equal numbers of pupils who do and who do not wish to study these subjects further (ratios of nearly 1:1).

On the contrary, foreign languages, chemistry, and biology in a ratio of about 1.5:1, and geometry with a ratio of 2.5:1, are the subjects in which the greater number of pupils have stated their opposition toward studying further in these fields.

[12] "Students dreading" in parentheses.

"Dreads".—The most unanimous results, however, were the statements of those students who had dreaded to begin their study of these subjects. Each subject revealed the greater numbers of students stating their indifference or opposition toward election of courses in these subject fields.

The intensity of their desire to elect or not to elect courses in these fields was indicated by smaller numbers of students and with differing proportions of "not-dreads"(36%–55%) and of "dreads"(23%–100%). Comparatively, in seven high school and three college subjects, the "dreads" described their desire as being strong to elect or not to elect courses in these fields in greater numbers, expressed in percentages, than did the "not-dreads".

TABLES IA, IB, IC, ID—STUDENT PERSONNEL

TABLE IA

CLASSIFICATION OF STUDENTS

Class	(A)[13]	(B)[13]	Total	%
Seniors	7	13	20	9
Juniors	35	82	117	52
Sophomores	69	2	71	32
Freshmen	14	0	14	6
Special	1	2	3	1
	126	99	225	100

[13] (A) and (B) are frequencies of students of educational psychology and of high school methods, respectively. This footnote applies to Tables IA, IB, IC, and ID.

CONCLUSIONS AND RECOMMENDATIONS

Significant for teacher guidance of learners in various subject areas are the findings: (1) that the development and persistence or change of attitudes accompanying achievement in these areas are in part dependent upon their earlier attitudes of dread or not-dread of experiences which these subject areas imply; and (2) that election of courses likewise is influenced by the fact of dread or not-dread of these subjects preceding study in these areas. Although these findings are limited by the relative amounts of the frequencies, for those of "dreads" are much smaller than those of "not-dreads", they reveal trends that are consistent and should be directive toward techniques which will guide pupils in their initial learnings in a given field. Hence the following recommendations are in order.

TABLE IB

STUDENT RATINGS OF THEIR ACHIEVEMENT LEVELS

Level	IN HIGH SCHOOL (A)	(B)	Total	%	Level	IN COLLEGE (A)	(B)	Total	%
A	65	58	123	56	A	7	7	14	6
B	56	36	92	42	B	75	58	133	61
C	2	3	5	2	C	37	32	69	32
					D	2	0	2	1
	123	97	220	100		121	97	218	100

TABLE IC

SIZE OF HIGH SCHOOL ATTENDED

School	(A)	(B)	Total	%
Large city	38	30	68	30
Small city	28	21	49	22
Town	53	46	99	44
Consolidated	6	1	7	3
Rural	1	0	1	0.4
	126	98	224	99.4

TABLE ID

RATINGS ON A. C. E. PSYCHOLOGICAL EXAMINATION

(For College Freshmen)

Rating	(A)	%
A (2, 1)[14]	14	11
B (4, 3)	39	31
C (6, 5)	41	33
D (9, 8, 7)	30	24
	124	99

[14] Scaled values of 1 to 9 are the means used to interpret the raw scores earned on this test.

SUGGESTIONS TO TEACHERS

1. *Teachers of algebra and geometry.*—Whereas "dreads" regard their own ability or inability as contributing causally to their difficulties in achieving satisfactorily in these subjects, and whereas "not-dreads" regard the teacher's guidance as the more important contributing factor, it is recommended that all "dreads" be discovered by the close of the

TABLE II

FREQUENCIES OF STUDENTS WHO HAD, OR HAD NOT, DREADED: (N=225)

HIGH SCHOOL SUBJECTS

SUBJECT	% REPORTING	% NOT DREADS	% DREADS
English	93	93	7
General science	50	89	11
Biology	43	86	14
World history	67	85	15
American history	87	82	18
Algebra	89	78	22
Foreign languages	86	77	23
Chemistry	28	69	31
Geometry	84	65	35

COLLEGE SUBJECTS

SUBJECT	% REPORTING	% NOT DREADS	% DREADS
English literature	55	94	6
English composition	98	79	21
American history	37	77	23
Foreign languages	48	76	24
Geography	28	75	25
Biology	36	65	35
Chemistry	29	45	55

TABLE III

NUMBERS OF STUDENTS RESPONDING TO QUESTIONS CONCERNING THE SUBJECTS

SUBJECTS	STATUS Initial L, I, D	STATUS Final L, I, D		DESIRE TO ELECT COURSE IN SAME SUBJECT			INTENSITY OF THIS DESIRE		
HIGH SCHOOL	(N)	Not Dr.	Dr.	(N)	Not Dr.	Dr.	(N)	Not Dr.	Dr.
English[15]	209	194,	15	204	194,	10	98	90,	8
History:									
American	196	160,	36	186	151,	35	80	63,	17
World	150	128,	22	151	129,	22	59	51,	8
Mathematics:									
Algebra	200	155,	45	197	155,	42	78	63,	15
Geometry	188	123,	65	187	124,	63	73	45,	28
Languages	193	148,	45	190	146,	44	95	70,	25
Science:									
General	113	100,	13	114	101,	13	48	45,	3
Biology	96	82,	14	97	86,	11	36	31,	5
Chemistry	62	43,	19	63	44,	19	29	19,	10
Geography[16]	20	18,	2	20	18,	2	8	7,	1
COLLEGE									
English:									
Composition	221	175,	46	199	158,	41	78	57,	21
Literature	124	117,	7	125	119,	6	71	65,	6
History:									
American	84	65,	19	81	62,	19	42	34,	8
Mathematics[16]	36	32,	4	35	31,	4	14	13,	1
Languages	107	81,	26	110	82,	28	56	40,	16
Science:									
Biology	82	53,	29	79	50,	29	35	26,	9
Chemistry	65	29,	36	65	31,	34	27	13,	14
Geography	60	45,	15	60	45,	15	29	23,	6

Legend: (English)

[15] The number of students (209) who indicated their liking (or indifference or dislike) of English during the earlier and latter parts of the semester of study included those (194) who had not dreaded and those (15) who had dreaded to study this subject.

Among those (204) who stated their desire to elect or not to elect another course in English, there were 194 students who had not and 10 who had dreaded to study English. The intensity of their desire was reported by 98 students, of whom there were 90 who had not, and eight who had, dreaded to study the subject.

[16] Subject was not included in this study, owing to its small frequency of cases.

TABLE IVA

RATIOS[17] OF FREQUENCIES OF STUDENT LIKE (L), INDIFFERENCE (I), OR DISLIKE (D) of SUBJECTS: (FINAL STATUS)

HIGH SCHOOL SUBJECTS

SUBJECT	RATIOS L	I	D
English[18]	20	5	1
American history	7	2	1
Not dreads	21	4	1
Dreads	*1*	1	1.5
World history	6	2	1
Algebra	5	2	1
Not dreads	13	4	1
Dreads	1	1	1+
Geometry	2	*1*	1
Not dreads	6	2	1
Dreads	1	*1*	2
Foreign languages	3	*1*	1
Not dreads	8	2	1
Dreads	*1*	2	4
General science	6	3	1
Biology	14	5	1
Chemistry	3	1	1

COLLEGE SUBJECTS

SUBJECT	RATIOS L	I	D
English composition	3	2	1
Not dreads	6	3	1
Dreads	1	2	2
English literature	6	*1*	1+
American history	5	2	1
Foreign languages	4	*1*	2
Not dreads	4	*1*	1
Dreads	2	*1*	3
Biology	3	*1*	1
Not dreads	4	*1*	1
Dreads	1	1	*1*
Chemistry	1+	1	*1*
Not dreads	3	2	1
Dreads	1	1	1+
Geography	3	*1*	2

[17] Ratios are expressed, to the nearest integer, in percentages. Frequencies pertain: (i) to all students; (ii) to "not dreads" and "dreads" when the minimum frequency of "dreads" is 25. *1* indicates the frequency, other than that of dislike, upon which the computations were based.

[18] Legend: English (high school)

Since the frequencies of L, I, D are 160, 41, and 8 respectively, there are, among all the students responding to the questionnaire in high school English, approximately 20 times as many "likes" as "dislikes" and 5 times as many "indifferents" as "dislikes".

TABLE IVB

RATIOS[19] OF FREQUENCIES OF STUDENT LIKE (L) AND NON-LIKE (I+D)

(Initial and Final Status)

HIGH SCHOOL SUBJECTS

SUBJECT	(N)	STATUS INITIAL	FINAL
English	209	4:1	3:1
American history	196	2:1	2:1
World history	150	2:1	2:1
Algebra	200	2:1	2:1
Geometry	188	*1*:*1*	1:1
Foreign languages	193	2:1	1:1
General science	113	2:1	2:1
Biology	96	3:1	2:1
Chemistry	62	1:1	1:1

COLLEGE SUBJECTS

SUBJECT	(N)	STATUS INITIAL	FINAL
English composition	221	1:1	1:1
English literature	124	5:1	3:1
American history	84	2:1	2:1
Foreign languages	107	3:1	2:1
Biology	82	1:1	1:1
Chemistry	65	2:1	2:1
	Reversed (NonL vs. L)		
Geography	62	2:1	1:1

[19] Ratios are expressed, to the nearest integer, in percentages. Frequencies for both initial and final status refer to the total number of students participating.

semester preceding their study of these subjects, and that instruction be patterned to decrease their difficulties and to guide them toward their most efficient mastery of these subjects.

2. *Teachers of all subjects.*—(a) Despite the relatively small number of "dreads" in certain subjects and the heavy preponderance of like versus indifference or dislike of subjects, procedures as suggested above would tend to reduce the negative and withdrawing attitudes of the "dreads", possibly to the extent of converting "dreads" into active, understanding, and eager searchers into the experiences thus opened to them, not only in this subject but in the more advanced and related subject areas.

(b) Whereas those who were indifferent to or disliked the subject tended to indicate the teacher as a contributing factor of greater significance than is their own ability or inability, and whereas those who liked the subjects regarded the nature of the subject as the factor of greatest influence, it would be helpful if teachers of these subjects would challenge their methods of teaching in the light of these findings. Whether the subject

be English, history, foreign languages, mathematics, or science on the high school or college level, it would seem wise for teachers of these subjects to analyze their techniques critically. It is suggested that they attack this problem by asking themselves such questions as: (1) Am I presenting these subject areas in logical ways? Or, am I guiding pupils in building more efficient habits of perceiving the data of a given problem precisely and comprehensively, of recalling facts pertinent to this problem, and especially of reacting toward it purposively? (2) Thus, have I directed their building of an hypothesis concerning the solution of this problem, and their testing of this hypothesis? (3) If the hypothesis does not work, do my pupils know what to do next, especially where to "begin again" in order to develop a second hypothesis and as many more "trial guesses" as prove to be necessary before the solution is found? (4) Will my pupils generalize this solution, and therefore increase the probability that they will recognize its reappearance as a factor in a similar situation and be able to act appropriately toward that situation?

Such constructive changes can become increasingly educative both for teachers who review and in part rebuild their methods of approach, of presentation, and of guidance of pupil activities, and also for pupils who can now approach initial learnings eagerly with increasing assurance that successes, not errors, are for them becoming realities.

TABLE Va[20]

VALUES OF X^2, WITH CORRESPONDING PROBABILITY (P)

	HIGH SCHOOL	X^2	(N)	COLLEGE	X^2
Part 1:	Foreign languages	45.9519	193		
	American history	39.7954	196		
	Geometry	29.1701	188		
	Algebra	27.2601	200		
			221	English composition	25.5538
	World history	17.2312	150		
	English	16.8195	209		
	General science	13.5108	113		
			107	Foreign languages	10.6460
	Biology	9.7847	96		
			84	American history	7.9585
			124	English literature	7.2201
Part 2:			82	Biology	5.2040
Part 3:			65	Chemistry	2.9989
			60	Geography	2.2222
	Chemistry	0.3321	62		

TABLE Vb[20]

(Based upon Table Va)

SUBJECTS IN:	n	X^2	
Part 1	1	>6.635	P<.01
Part 2	1	5.2039	P<.05
Part 3	1	0.3321—2.9988	P>.05 or (.10—.70)

[20] Tables Va and Vb: See the text of this article for the meaning and use of these tables. Meaning of:

(N) is the number of students responding to a subject questionnaire.

"n" is the "number of degrees of freedom" in the 2 by 2, or fourfold table, upon which the computations were based.

X^2: The table of X^2 is presented in the text: R. A. Fisher, *Statistical Methods for Research Workers*, pp. 118-19.

P (Probability): *ibid.*, pp. 82, 83.

TABLE VI[21]

STUDENT DESIRE TOWARD FURTHER STUDY IN THESE SUBJECT AREAS

ALL STUDENTS

HIGH SCHOOL Subject	Ratios Y>N	Y=N	N>Y	COLLEGE Subject	Ratios Y>N	Y=N	N>Y
English	1.9			English composition		1.1	
American history[22]		1.0		English literature	2.6		
World history		1.1		American history		1.0	
Algebra			1.7	Foreign languages		1.1	
Geometry			2.5	Biology			2.4
Foreign languages			1.4	Chemistry			1.8
General science		1.15		Geography			2.8
Biology		1.0					
Chemistry			1.6				

STUDENTS NOT DREADING } THESE SUBJECTS
STUDENTS DREADING[23] }

HIGH SCHOOL Subject	Ratios Y>N	Y=N	N>Y	COLLEGE Subject	Ratios Y>N	Y=N	N>Y
English	2.0		(2.3)	English composition	1.3		(5.8)
American history	1.3		(2.9)	English literature	3.0		(5.0)
World history		1.15	(6.3)	American history	1.3		(2.1)
Algebra		1.9	(6.0)	Foreign languages	1.5		(6.0)
Geometry		1.6	(9.5)	Biology		1.5	(8.6)
Foreign languages		1.0	(6.3)	Chemistry		1.1	(3.9)
General science		1.0	(5.5)	Geography			2.0 (14.1)
Biology		1.15	(2.7)				
Chemistry			1.3 (2.8)				

[21] Expressed in ratios of frequencies of those who:
(Y) wish to elect the subject,
(N) wish not to elect subject.

[22] Legend: (American history)
Among all of the pupils responding to this question concerning the study of American history in high school, there are approximately equal numbers of pupils who do and who do not wish to study this subject further. Their desire is expressed by the ratio of 1:1.
Among those pupils who had not dreaded their study of this subject, the larger number of pupils have indicated their desire (expressed by the ratio of 1.3:1) to study another course in the same subject.
Among those pupils who had dreaded their study of this subject, the greater number of pupils indicated their opposition (revealed by the ratio of 2.9:1) toward further study in this field.

[23] "Students dreading" in parentheses.

NOV 24 1944

SCHOOL OF EDUCATION
LIBRARY

The Journal of Experimental Education

A periodical report of scientific investigations relating to child development, curriculum, learning, teaching, supervision, measurements, statistics, and experimental techniques.

Volume XII | DECEMBER, 1943 | Number 2

CHILD DEVELOPMENT

CONTENTS

PAGE

Teacher Child Contacts in Nursery Schools: *Catherine Landreth, Gladys M. Gardner, Bettie C. Eckhardt and Anne D. Prugh* ---------- 65

A Test of General Information for Children of Preschool Age: *Janet Smith* --- 92

The Effects of Early Institutional Care on Adolescent Personality: *William Goldfarb* ---------- 106

Child Development and the Curriculum: Some General Principles: *Arthur T. Jersild and Charlotte Fehlman* ---------- 130

$5.00 A YEAR | PUBLISHED QUARTERLY | $1.50 A COPY

Edited and Published by A. S. Barr, Professor of Education, University of Wisconsin, Madison, Wisconsin.

Entered as second-class matter October 17, 1938 at the post office at Madison, Wisconsin under the act of March 3, 1879.

EDITORIAL BOARD

A. S. Barr, Chairman, Professor of Education, University of Wisconsin, Madison, Wis.

Carter V. Good, Professor of Education, University of Cincinnati, Cincinnati, Ohio. Editorially responsible for materials on learning, teaching and supervision, published each September.

J. Wayne Wrightstone, Assistant Director, Bureau of Reference, Research and Statistics, Board of Education of the City of New York, 110 Livingston Street, Brooklyn, New York. Editorially responsible for materials on curriculum construction, published each June.

Edward E. Cureton, Senior Personnel Technician, Personnel Research and Occupational Analysis Section, The Adjutant General's Office, 270 Madison Avenue, New York 16, N. Y. Editorially responsible for materials on measurements, statistics, and methods of experimental research, published each March.

Arthur T. Jersild, Professor of Education, Teachers College, Columbia University, New York City. Editorially responsible for materials on child welfare, guidance, and development, published each December.

CONTRIBUTING EDITORS

Gilbert L. Betts, Supervisor of Graduate Research in Education, Colorado State College, Fort Collins, Colorado.

Leo J. Brueckner, Professor of Education, University of Minnesota, Minneapolis, Minnesota.

Oscar K. Buros, Associate Professor of Education, Rutgers University, New Brunswick, New Jersey.

Otis W. Caldwell, General Secretary, The American Association for the Advancement of Science, Boyce Thompson Institute for Plant Research, Yonkers, New York.

Leslie L. Chisholm, Associate Professor of Education, State College of Washington, Pullman, Washington.

Herbert S. Conrad, College Entrance Examination Board, Princeton, N. J.

Stephen M. Corey, Professor of Educational Psychology, University of Chicago, Chicago, Illinois.

Robert A. Davis, Professor of Education, Director of Bureau of Educational Research, University of Colorado, Boulder, Colorado.

Harl R. Douglass, Director of College of Education, University of Colorado, Boulder, Colorado.

Jack W. Dunlap, Associate Professor of Educational Psychology, University of Rochester, Rochester, New York.

Harold A. Edgerton, Assistant Professor of Psychology, Ohio State University, Columbus, Ohio.

Lt. Comdr., A. C. Eurich, USNR, Officer in charge, Standards and Curriculum, Bureau of Naval Personnel, USN, 3130 Wisconsin Avenue, N. W., Washington 15, D. C.

Lt. Col. John C. Flanagen, Chief, Psychological Branch, Office of Air Surgeon, Headquarters, Army Air Forces, Washington 25, D. C.

Harold E. Jones, Professor of Psychology and Director, Institute of Child Welfare, University of California, Berkeley 4, California.

Noel Keys, Professor of Education and Lecturer in Human Relations, University of California, Berkeley, California.

Edward A. Lincoln, Consulting Psychologist, Halifax, Massachusetts.

T. E. Newland, Lt., USNR, 2702 Wisconsin Avenue, N. W., Washington 7, D. C.

C. W. Odell, Associate Professor of Education, University of Illinois, Urbana, Illinois.

Willard C. Olson, Professor of Education, Director of Research in Child Development, University of Michigan, Ann Arbor, Michigan.

W. E. Peik, Dean and Professor of Education, University of Minnesota, Minneapolis, Minnesota.

S. L. Pressey, Professor of Educational Psychology, Ohio State University, Columbus, Ohio.

Clarence E. Ragsdale, Professor of Education, University of Wisconsin, Madison, Wisconsin.

H. H. Remmers, Director, Division of Educational Reference, Professor of Education and Psychology, Purdue University, Lafayette, Indiana.

Henry D. Rinsland, Professor of Education and Director of Educational Research, The University of Oklahoma, Norman, Oklahoma.

Robert T. Rock, Jr., Professor of Psychology, Head of Dept. of Psychology, Graduate School, Fordham University, New York City.

Phillip J. Rulon, Associate Professor of Education, Harvard University, Cambridge 38, Massachusetts.

Douglas E. Scates, Associate Professor of Education, Duke University, Durham, North Carolina.

David Segel, Educational Consultant, Specialist in Tests and Measurements, Federal Security Agency, U. S. Office of Education, Washington, D. C.

Paul W. Terry, Professor of Educational Psychology, University of Alabama, University, Alabama.

Helen Thompson, Clinic of Child Development, Research Associate, Yale University, New Haven, Connecticut.

Robert L. Thorndike, Associate Professor of Education, Teachers College, Columbia University, New York City.

Herbert A. Toops, Professor of Psychology, Ohio State University, Columbus, Ohio.

T. L. Torgerson, Professor of Education, University of Wisconsin, Madison, Wisconsin.

Helen M. Walker, Professor of Education, Teachers College, Columbia University, New York City.

Beth L. Wellman, Professor of Psychology, Child Welfare Research Station, State University of Iowa, Iowa City, Iowa.

Guy M. Wilson, Professor of Education, Boston University, Boston, Massachusetts.

Paul A. Witty, Professor of Education, Director of Psycho-Educational Clinic, School of Education, Northwestern University, Evanston, Illinois.

Ernest R. Wood, Professor of Education, New York University, New York City.

DEMOCRAT PRINTING COMPANY
MADISON, WISCONSIN

Property of the Library
Texas Tech University